

FACULTÉ DES LETTRES DE L'UNIVERSITÉ DE PARIS

Remarques sur la formation
de la
Théorie abstraite des ensembles

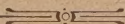
Étude historique et critique

THÈSE COMPLÉMENTAIRE POUR LE DOCTORAT ÈS LETTRES
(PHILOSOPHIE)

PRÉSENTÉE

par **J. CAVAILLÈS**

Ancien Elève de l'Ecole Normale Supérieure
Agrégré de Philosophie



PARIS

HERMANN & C^{ie}, ÉDITEURS

6, Rue de la Sorbonne

1937

THÈSE
POUR
LE DOCTORAT



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41551601_0014

FACULTÉ DES LETTRES DE L'UNIVERSITÉ DE PARIS

Remarques sur la formation
de la
Théorie abstraite des ensembles

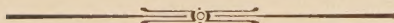
Étude historique et critique

THÈSE COMPLÉMENTAIRE POUR LE DOCTORAT ÈS LETTRES
(PHILOSOPHIE)

PRÉSENTÉE

par **J. CAVAILLÈS**

Ancien Elève de l'Ecole Normale Supérieure
Agrégré de Philosophie



PARIS
HERMANN & C^{ie}, ÉDITEURS
6, Rue de la Sorbonne

1937

La Faculté n'entend donner aucune approbation ni improbation aux opinions émises dans les Thèses; ces opinions doivent être considérées comme propres à leurs auteurs.





INTRODUCTION

IL est peu d'histoires aussi traversées que celle de la théorie des ensembles. Après un départ difficile, dans une méfiance justifiée par l'échec des spéculations antérieures sur l'infini actuel ⁽¹⁾, à peine la construction terminée reçoit-elle la consécration du monde mathématique, que la découverte des antinomies vient en secouer les bases d'un ébranlement qui menace de s'étendre à l'analyse, contre lequel aucun remède ne s'est jusqu'ici révélé efficace. En 1925, VON NEUMANN concluait un nouvel essai d'axiomatisation par la mise en évidence du caractère irréductiblement non catégorique ⁽²⁾ de tout système d'axiomes qui voudrait sauvegarder l'essentiel de la théorie : « nous ne pouvons faire autre chose que constater qu'ici réside une nouvelle objection contre la théorie des ensembles et que pour l'instant aucun moyen de réhabilitation n'est en vue » ⁽³⁾. Aussi bien ces entreprises ont-elles eu pour résultat de provoquer une juxtaposition de théories plus ou moins différentes de l'œuvre cantorienne primitive :

(1) Par exemple la *Géométrie de l'Infini* par FONTENELLE, tissu de raisonnements incorrects.

(2) C'est-à-dire que le système d'axiomes considéré comme définition implicite ne détermine pas de façon univoque son objet. Nous renvoyons pour cette notion à J. CAVAILLÈS, *Méthode axiomatique et formalisme*. Paris, Hermann, 1937, ch. II.

(3) VON NEUMANN, *Eine Axiomatisierung der Mengenlehre*. *Journ. f. Math.*, t. 154, p. 240.

il n'y aurait plus une théorie des ensembles, mais celle de ZERMELO-FRAENKEL, celle de VON NEUMANN, celle de BROUWER. Or, les utilisations, au moins dans les mathématiques classiques, exigent l'unité. Ici apparaît un autre fait qui n'est pas le moindre paradoxe de cette histoire, les recours toujours plus nombreux à une théorie contestée dans les mathématiques incontestables. Dès 1873, dans une lettre à DUBOIS-REYMOND, WEIERSTRASS se servait de la dénombrabilité des nombres algébriques, — que CANTOR venait de démontrer, — pour étendre un théorème sur les fonctions de variables réelles ⁽¹⁾. A partir de 1884 avec les mémoires de POINCARÉ sur la structure du domaine d'existence des fonctions automorphes, de MITTAG-LEFFLER sur la possibilité de construction de fonctions avec des singularités données, se développent une série d'applications toujours plus importantes en analyse ⁽²⁾. Contrairement à l'attente de CANTOR, tourné vers les fonctions analytiques, c'est surtout à l'étude des fonctions de variables réelles que la théorie fait faire les progrès les plus considérables. Les travaux de BAIRE, BOREL, LEBESGUE donnent des modes nouveaux d'engendrement ou de mise en correspondance dans l'espace cartésien et transforment la théorie de l'intégration. La crise théorique déclenchée en 1905 ne fait qu'embarrasser de restrictions et de scrupules les raisonnements sans arrêter leur puissance interne d'expansion. Aujourd'hui restrictions et scrupules passent au second plan quoique le problème du fondement semble moins près que jamais d'une solution, et trois disciplines manifestent une dépendance directe et essentielle par rapport à la théorie des ensembles : l'analyse fonctionnelle ou topologie générale fondée en 1909 par FRÉCHET, qui elle-

(1) Lettre du 15 décembre 1874, *Acta Mathematica*, t. 39 (1924), p. 206.

(2) Cf. A. DENJOY, *Introduction à la théorie des fonctions de variables réelles*. Paris, Hermann, 1937, p. 12 : « La plus grande révolution que la mathématique ait connue depuis l'invention du calcul différentiel et intégral venait de s'ouvrir. »

même trouve chaque jour de nouveaux champs d'application, la théorie de la mesure et de l'intégration développée par BOREL, LEBESGUE et DENJOY, enfin la topologie combinatoire. Sans doute ne s'agit-il que d'utilisations chaque fois nettement délimitables, et certains mathématiciens ont-ils même soin de séparer les résultats obtenus à l'aide des parties les plus contestées de la théorie, par exemple les différents énoncés plus ou moins forts de l'axiome de choix. Mais, une méthode n'est pas chose isolable arbitrairement pour les besoins d'un problème. L'exigence même de sécurité oblige à la dépouiller de son vêtement accidentel, pris dans un cas particulier, à préciser les conditions nécessaires et suffisantes de son application. Si elle réussit avec certains ensembles, ce ne peut être qu'en vertu de propriétés définissables en eux : il n'y aura mathématique rigoureuse que lorsqu'on aura, par l'investigation même des procédés, défini le champ d'objets auxquels ils conviennent. D'autre part, ces procédés s'enchaînent dans un complexe organique : c'est un état d'esprit qui constitue la base secrète d'où ils jaillissent avec une apparente imprévisibilité, à quoi il faut revenir si l'on veut vraiment comprendre et non plus exposer mais continuer. Parenté profonde qui unit — sans qu'ils en aient même pleine conscience, comme une entente implicite — les chercheurs d'une même discipline au moins dans une période donnée, *vis a tergo*, élan objectif de la recherche elle-même. Cette double unité, justification, pour une théorie, de son existence mathématique, la théorie des ensembles la possède-t-elle ? Rechercher la première est travail technique : on a vu à quelles difficultés il se heurte. Mais peut-être, ne fut-ce que pour évaluer ses chances d'aboutir, est-il possible d'essayer de déceler la seconde là où elle existe, sous les juxtapositions, peut-être factices, d'une construction arbitraire. C'est ici que l'histoire intervient.

L'histoire mathématique semble, de toutes les histoires, la moins liée à ce dont elle est véhicule ; s'il y a lien, c'est

a parte post, servant uniquement pour la curiosité, non pour l'intelligence du résultat : l'après explique l'avant. Le mathématicien n'a pas besoin de connaître le passé, parce que c'est sa vocation de le refuser : dans la mesure où il ne se plie pas à ce qui semble aller de soi par le fait qu'il est, dans la mesure où il rejette autorité de tradition, méconnaît un climat intellectuel, dans cette mesure seule il est mathématicien, c'est-à-dire révélateur de nécessités. Cependant avec quels moyens opère-t-il ? L'œuvre négatrice d'histoire s'accomplit dans l'histoire. Double liaison : avec les problèmes posés et étudiés dans un temps — choix de la rébellion —, avec les méthodes déjà existantes, matière où forger le nouvel instrument. Dans les deux cas, l'arbitraire individuel ou le style d'un milieu ne suffisent pas à expliquer : même si l'on concevait les mathématiques comme système en soi, les sinuosités du processus de révélation seraient en relation avec la structure des parties révélées. Autrement dit, il y a une objectivité, fondée mathématiquement, du devenir mathématique ; c'est l'exigence d'un problème qui oblige à dépouiller une méthode d'accidents qu'aucune réflexion n'apercevait inutiles, c'est la vigueur interne d'une méthode qui dépasse son champ primitif d'application et pose de nouveaux problèmes. Connexion du système mathématique d'un temps, pour lequel la réciprocité d'actions exclut aussi bien les lacunes intérieures, que l'imagination d'un vide extérieur où se situerait la fixité d'un modèle.

Dans cette indétermination, le départ est difficile à faire entre enchaînements psychologiques ou sociaux et nécessité mathématique — on le voit à propos de ces découvertes simultanées, explicables à la fois par la présence de l'instrument ou l'acuité d'une question, et parce que l'attention y était socialement portée. Les éléments indubitablement objectifs sont ces nœuds que constituent les théories : ici dans l'enchevêtrement entre notions et méthodes disparaissent les liens de causalité au profit des relations d'in-

telligibilité. Processus dont l'exposé systématique — axiomatisation par exemple — est manifestation mais non indispensable. Il est possible d'apercevoir l'unité interne dans l'intuition centrale d'où procèdent les enchaînements, de la déceler aussi méthodiquement par la seule description du phénomène. Suivre la genèse des notions, préciser surtout leurs liens effectifs avec les problèmes et isoler les procédés généraux dont le geste intérieur rapproche le plus de l'intuition centrale impossible à décrire, tel est le travail — à la fois soumis à l'histoire et critique de celle-ci au nom de ses résultats — qui a quelque chance d'aboutir à un résultat objectif. Devant cet effort pour situer au sein du système mathématique, en même temps que caractériser une structure autonome, devraient disparaître les architectures factices, peut-être même aussi se deviner la stérilité de certaines questions. C'est un étonnement toujours renouvelé que les inflexions inattendues du devenir mathématique, son ironique abandon des voies qu'un essai de prévision systématique ouvrait devant lui. A ce problème d'une explication de l'inégale fécondité des méthodes, une étude critique suffisamment ample de théories voisines pourrait peut-être apporter quelque lumière. Mais il ne s'agit ici que d'une application limitée du procédé à la question spéciale posée à propos de la théorie des ensembles : apparition nécessaire ou, et dans quelle mesure, fantaisie historique, structure autonome ou pluralité enrobée dans un système contingent.

CHAPITRE I

PRÉHISTOIRE

NOMBRES IRRATIONNELS. SÉRIES
TRIGONOMÉTRIQUES.
THÉORIE DE LA CROISSANCE

Le XIX^e siècle mathématique s'ouvre par une crise de l'infini : aux premières objections contre infiniment grand et infiniment petit, le XVIII^e avait fait succéder une confiance aveugle, pour le passage à la limite, par exemple, et l'emploi des séries. En 1826, ABEL pouvait se plaindre encore qu'on raisonne sur des séries divergentes : « elles sont quelque chose de bien fatal et c'est une honte qu'on ose y fonder aucune démonstration... Ce sont elles qui ont fait tant de malheurs et causé tant de paradoxes (1) ». Des malheurs et des paradoxes la méthode d'exhaustion avait su se garder tant que réservée à la géométrie, garantie contre les absurdités par l'intuition sensible. La révolution cartésienne, en assurant un parallélisme complet entre nombres et étendue, en inaugurant l'étude des variations de fonctions, incorporait le continu à l'algèbre : celle-ci, devenue autonome, devait définir par ses propres moyens les objets employés, préciser, étendre elle-même les modes d'emploi. D'où deux sortes de problèmes : donner un fondement purement ana-

(1) ABEL, *Untersuchungen über die Reihe* $1 + \frac{m}{1}x + \frac{m \cdot (m-1)}{1 \cdot 2}x^2 + \dots$
Journal de Crelle, t. 1 (1826), p. 311.

lytique aux notions et propriétés assurées jusque-là seulement par l'évidence de leur corrélat géométrique ; développer les unes et les autres hors des limites arbitraires imposées par l'intuition. Etablissement rigoureux de la théorie des fonctions continues, extension des opérations analytiques, double travail du XIX^e siècle qui exigeait, pour remplacer l'intuition défailante, la création des notions essentielles de la théorie des ensembles.

En 1799, BOLZANO examine ⁽¹⁾ les différentes démonstrations du théorème que « toute fonction algébrique rationnelle d'une variable peut être décomposée en facteurs réels du 1^{er} ou du 2^e degré » ; aucun raisonnement employé n'échappe à l'alternative : cercle vicieux, recours à l'intuition géométrique. Ainsi ce postulat posé au préalable (et d'où le théorème se déduit aisément) « qu'une ligne continue à courbure simple dont les coordonnées sont d'abord positives puis négatives doit rencontrer l'axe des abscisses au moins en un point situé entre ces ordonnées ». Rien ne manque à son évidence sensible ; si on veut rester dans l'analyse, il faut traduire — toute fonction continue de x qui dans un intervalle prend des valeurs positives et négatives doit au moins une fois s'annuler — et « chercher un fondement objectif ». La définition de la continuité est la définition moderne. BOLZANO utilise un lemme important aussi bien par son énoncé que par son mode de démonstration. Soit un caractère M que possèdent une partie des nombres ⁽²⁾ de l'intervalle $(0 \dots 1)$; en particulier tous les nombres inférieurs à u . On peut déterminer dans $(0 \dots 1)$ un nombre U qui soit le plus grand des nombres tels que

(1) BOLZANO (I), p. 4-14. Les numéros renvoient à la bibliographie méthodique.

(2) Nous employons dès maintenant le mot nombre pour désigner un élément quelconque de l'intervalle arithmétique $(0 \dots 1)$. Pour BOLZANO et ses contemporains, c'était encore le mot « valeur » ou « grandeur » qui s'imposait, tant était encore géométrique la notion de continu.

u . Soit, en effet, $u + v$ un des nombres (existant par hypothèse) qui ne possèdent pas la propriété M ; on examinera si tous les nombres entre u et $u + \frac{v}{2}$ possèdent M ; en cas de réponse négative, si tous entre u et $u + \frac{v}{2^2}$, etc. Deux éventualités : ou on arrive à une réponse affirmative pour $u + \frac{v}{2^m}$, ou l'on a toujours des réponses négatives. Dans le 2^e cas $U = u$; dans le 1^{er} cas, on recommence le même raisonnement pour l'intervalle :

$$\left(u + \frac{v}{2^m} \dots u + \frac{v}{2^m} + \frac{v}{2^{m+n}} \right)$$

S'il est possible de poursuivre ainsi indéfiniment dans des intervalles emboîtés, le nombre U sera défini comme la limite de la suite convergente.

$$u, u + \frac{v}{2^m}, u + \frac{v}{2^m} + \frac{v}{2^{m+n}}, \dots$$

dans le cas contraire, il coïncide avec un des termes de la suite.

CANTOR (1) donne une analyse générale de ce procédé pour étudier une propriété *distributive*, c'est-à-dire telle que si elle appartient aux éléments d'un intervalle (ou à un ensemble contenu dans cet intervalle), elle appartient à ceux d'au moins une portion de cet intervalle divisé en un nombre fini de portions. Cette polychotomie finie qui se poursuit de proche en proche dans des intervalles emboîtés de plus en plus petits, est « une méthode qui peut être remplacée difficilement par une autre ; l'essentiel en est très vieux, et, dans les temps modernes, utilisé par LAGRANGE, LEGENDRE, DIRICHLET, CAUCHY » (2). On verra l'usage qu'en

(1) CANTOR (XV), p. 212. La pagination est, de même que pour les mémoires de DEDEKIND, empruntée aux recueils CANTOR (XXVIII) et DEDEKIND (V),

(2) *Ibid.*

feront non seulement CANTOR lui-même, mais PEANO et, en particulier, BOREL et LEBESGUE pour la démonstration de leur lemme. Si BOLZANO n'est pas ici novateur, du moins l'application est-elle originale : les nombres de l'intervalle sont des éléments d'ensemble pour lesquels on ne suppose rien sinon qu'ils sont ordonnés et qu'on peut décider s'ils possèdent ou non la propriété M .

Le théorème principal peut être alors établi : soit a un nombre rendant la fonction $f(x)$ négative ; puisque $f(x)$ est continue, il existe un nombre η tel que :

$$\text{pour } \varepsilon < \eta \quad f(a + \varepsilon) < 0$$

(caractère M du lemme pour les nombres ε) ; soit U le nombre du lemme : on ne peut avoir ni

$$f(a + U) < 0$$

puisque en vertu de la continuité $f(x + U + \varepsilon)$ serait encore négatif pour ε assez petit, contrairement à la définition de U , ni

$$f(a + U) > 0$$

pour les mêmes raisons relativement à $f(a + U - \varepsilon)$. Donc

$$f(a + U) = 0.$$

Mais une rigueur complète n'est toutefois pas atteinte : il faut admettre que la suite convergente :

$$u, u + \frac{v}{2^m}, u + \frac{v}{2^m} + \frac{v}{2^{m+n}}, \dots$$

donne l'existence de U . BOLZANO croit le démontrer en prouvant que, la suite étant donnée, on peut calculer sa limite avec une approximation arbitraire, comme si la suite des approximations n'était pas à son tour une suite convergente. Ici encore subsistent donc irréductiblement cercle vicieux ou recours à l'intuition : seule la définition complète du nombre réel permet de les éviter.

* * *

Il est assez remarquable qu'en 1876 LIPSCHITZ — qui devait d'ailleurs adopter en 1877 une définition voisine pour son traité d'analyse, le « premier rigoureux » au dire de PASCH (1) — refuse encore toute originalité à la théorie de DEDEKIND dans *Stetigkeit und irrationalen Zahlen* : « je ne nie pas la rectitude de votre définition, mais j'estime qu'elle ne se distingue que dans la forme de l'expression, et non pour le fond, de celle que les anciens ont posée... Aussi souhaiterais-je que vous laissiez l'affirmation que des propositions comme $\sqrt{2} \cdot \sqrt{3} = \sqrt{6}$ n'ont jusqu'ici pas été démontrées. Je crois, en effet, qu'en particulier les lecteurs français seront convaincus avec moi que le ... [Ve] livre d'EUCLIDE contient les principes nécessaires et suffisants pour la démonstration de ce théorème » (2). Si grande était la difficulté de distinguer à ce degré de généralité entre évidence d'un acte et démonstration d'une possibilité dans un domaine abstrait. « EUCLIDE, répond DEDEKIND, peut appliquer sa définition de l'égalité des rapports à toutes les grandeurs... dont l'existence est admise pour de bonnes raisons. Mais pour fonder l'arithmétique sur la notion de rapport de grandeurs (ce qui n'a jamais été le but d'EUCLIDE) elle ne suffit absolument pas (3). » En réalité, le processus d'abstraction nécessaire se heurtait à la représentation traditionnelle des deux actes générateurs des objets mathématiques : compter et mesurer. ILLIGENS reprochait en 1889 dans les *Mathematische Annalen* (4) à la définition de CANTOR de ne concerner que des signes abstraits qui, n'étant ni multiplicités ni grandeurs, n'avaient rien de comparable

(1) PASCH (I), p. 1.

(2) *Correspondance Lipschitz Dedekind* in DEDEKIND (V), t. III, p. 469.

(3) *Ibid.*, p. 477.

(4) ILLIGENS (I).

aux nombres véritables. Pourtant, l'introduction des nombres complexes avait déjà nécessité l'abandon du « concept obscur de grandeur ». De HAMILTON à GRASSMANN le calcul symbolique rendait familières ces définitions, par leurs propriétés formelles, d'opérations effectuelles sur un matériel abstrait quelconque. Le recours à la notion d'ensemble est inévitable : c'est d'elle que doit partir une mathématique qui ne veut plus reconnaître l'absolu d'un acte. La quasi simultanéité des différentes définitions du continu numérique entre 1870 et 1880 est, à cet égard, suffisamment instructive.

Les trois méthodes de DEDEKIND-TANNERY ⁽¹⁾, de WEIERSTRASS et de HEINE-CANTOR ne se distinguent guère que par la présentation : DEDEKIND procède de façon plus descriptive en caractérisant par ses propriétés l'ensemble des nombres réels, WEIERSTRASS analyse la notion de nombre pour y trouver motif et possibilité de l'extension cherchée, enfin HEINE et CANTOR se préoccupent davantage des procédés de calcul. Dans tous les cas on « prend comme seule base, sur laquelle on doit s'être mis d'accord, l'arithmétique des nombres rationnels. » ⁽²⁾

La notion fondamentale pour DEDEKIND est celle d'extensibilité ; les nombres rationnels constituent un *corps* (système invariant par rapport aux 4 opérations) soumis à un ordre que caractérisent trois propriétés : 1° la définition habituelle de l'ordre (relation existant entre 2 éléments quelconques, asymétrique, non réfléchie, transitive) ; 2° entre deux éléments quelconques l'ordre en situe une infinité d'autres ; 3° tout élément sépare le système en deux classes non vides telles que tout élément de la première précède dans l'ordre tout élément de la deuxième (propriété distin-

(1) Nous n'exposons pas la méthode de Tannery parce qu'identique et — quoique trouvée de façon indépendante — postérieure à celle de Dedekind.

(2) L. de DEDEKIND à LIPSCHITZ, 1876, *l. c.*, p. 470.

quant l'ordre linéaire d'un ordre circulaire, ou plan). L'ensemble des nombres réels sera le corps maximum possédant ces trois propriétés : son inextensibilité s'appellera *continuité*. Pour l'obtenir il suffit d'exiger par rapport à l'opération impliquée dans la 3^e propriété la même fermeture que par rapport aux 4 opérations arithmétiques : toute répartition en deux classes soumises aux conditions précédentes (ou coupure) est déterminée par un élément de l'ensemble (dernier élément de la 1^{re} classe ou 1^{er} élément de la 2^e) : on voit qu'il est alors impossible d'ajouter un nouvel élément (1). L'ensemble L des nombres rationnels n'est pas inextensible : la répartition

$$x^2 < D < (x + 1)^2,$$

D n'étant pas un carré, est une coupure dont les deux classes n'ont ni premier ni dernier élément. Considérons l'ensemble Λ des coupures de l'ensemble L ; on peut y établir un ordre : une coupure en précède une autre si la 1^{re} classe de la 2^e contient des éléments de la 2^e classe de la 1^{re}. On démontre (2) que cet ordre possède les trois propriétés énoncées plus haut. De plus : tout élément a de L possède un représentant α dans Λ en vertu de la propriété 3^o de L (l'ordre est évidemment conservé) ; entre deux éléments quelconques de Λ il y a une infinité d'éléments α représentants d'éléments de L (si, en effet, dans L , deux classes correspondant à deux coupures diffèrent par plus d'un élément, elles diffèrent par une infinité d'éléments en vertu de la propriété 2^o). Enfin apparaît pour Λ une 4^e propriété : toute

(1) En effet son insertion dans l'ordre déjà établi provoquerait une coupure dans le système primitif : il coïnciderait donc, en vertu de la propriété 2^o, avec l'élément déterminant la coupure.

(2) La propriété 1^o est évidemment présente ; entre 2 coupures de L il y a une infinité de coupures, celles que déterminent les éléments de L ; enfin une coupure de L répartit bien les autres en deux classes, celles qui la précèdent et celles qui la suivent ; chacune des classes comportant évidemment des éléments (les coupures déterminées par les éléments de L en particulier).

coupure Π est déterminée de façon univoque par un élément et un seul (premier élément de la classe Π_2 ou dernier élément de la classe Π_1). En effet, la coupure Π définit une coupure P dans L (les deux classes sont formées respectivement des éléments dont les représentants sont dans Π_1 et Π_2) donc un élément Π de Λ : Π ne peut être que le 1^{er} élément de Π_2 ou le dernier de Π_1 , sans quoi il serait précédé dans Π_2 , ou suivi dans Π_1 d'éléments représentants d'éléments de L , ce qui est impossible. L'ensemble Λ est un corps si on définit les 4 opérations arithmétiques à partir des mêmes opérations effectuées sur les nombres rationnels des classes correspondant aux coupures (la somme de deux coupures est la coupure dont la 1^{re} classe est formée de tous les nombres rationnels inférieurs à la somme de deux nombres quelconques appartenant respectivement aux premières classes des deux sommants). Une suite convergente détermine bien ici un nombre fixe qui diffère aussi peu qu'on veut du terme général de u_n de la suite pour n assez grand : la répartition des nombres de Λ en classe de ceux pour lesquels, quel que soit N , on peut trouver un $n > N$ tel que u_n les dépasse, et classe des autres, constitue une coupure ; le nombre séparatif est la limite cherchée.

Le passage du système L au système Λ est produit par la « puissance créatrice » (1) de notre esprit : autrement dit, d'une part les éléments de Λ possèdent une existence aussi authentique (et de même ordre) que les nombres rationnels, d'autre part ils s'en distinguent justement par l'étape du processus créateur qui fait passer du nombre à la coupure (si elle est déterminée par un nombre rationnel) puis de la coupure au nombre (réel). Le modèle — et sans doute l'idée première — de semblables créations se trouve dans les idéaux de KUMMER dont l'introduction, « conduite avec assez de rigueur, doit être considérée comme absolument légi-

(1) Dedekind l. à Weber, 24 janvier 1888, in DEDEKIND (V), t. III, p. 489.

time » (1). C'est en acte la méthode axiomatique dans une de ses premières utilisations : « une infaillible méthode pour l'analyse [de toutes les suppositions explicites ou tacites] ... est de remplacer toutes les expressions nouvelles par des mots arbitraires dépourvus de sens, l'édifice doit, s'il est bien construit, ne pas en être ébranlé, et j'affirme que par exemple ma théorie des nombres réels soutient l'épreuve. » (2)

Au contraire, WEIERSTRASS (3) se place au point de vue génétique : « Le nombre, dans le cas le plus simple, est la représentation synthétique de l'adjonction de l'unité à l'unité. » De toute façon, il se rapporte à une multiplicité : « un nombre complexe est la représentation synthétique de groupes différents d'éléments homogènes entre eux. » (4) Ainsi le nombre résulte, de toutes façons, d'un processus d'abstraction appliqué à une collectivité : il y faut distinguer les éléments qui peuvent appartenir à des espèces différentes et leur agrégat (ANZAHL) fini ou infini (5). Opérations et relations quoique spécialisées pour chaque espèce, sont soumises à des principes généraux, où l'on peut trouver à peu près exactement les définitions correspondantes formulées plus tard par CANTOR pour la théorie des ensembles. La somme de 2 nombres est l'agrégat formé de la réunion de leurs éléments, le produit de α par β , l'agrégat obtenu en remplaçant chaque élément de β par un agrégat équivalent à α . On peut, dans un nombre, changer les éléments fondamentaux, sans changer le nombre : ainsi deux nombres sont égaux si, par un changement convenable, on

(1) *Ibid.*, p. 490.

(2) DEDEKIND l. à LIPSCHITZ, 27 juillet 1876, *ibid.*, p. 479.

(3) La théorie de WEIERSTRASS a été exposée en 1872 dans KOSSAK (I).

(4) *Ibid.*, p. 16.

(5) Lorsqu'un nombre est représenté par un agrégat fini, avec comme éléments des nombres rationnels, par exemple, il coïncide avec la somme de ses éléments.

peut les amener à constituer un agrégat identique des mêmes éléments. Mais il y a une définition plus générale de l'égalité s'appliquant aussi aux agrégats infinis : deux nombres sont égaux si tout nombre formé d'un agrégat fini d'éléments de l'un d'eux est partie intégrante de l'autre. Un nombre a dénotant un agrégat infini est fini s'il existe un nombre b formé d'un agrégat fini tel que tout nombre formé d'éléments de a (en agrégat fini) est contenu dans lui, la réciproque étant impossible (ce qui donne, d'après les définitions précédentes, $a < b$). Tel est le nombre irrationnel agrégat infini d'éléments qui sont des nombres rationnels (que l'on suppose déjà définis). KOSSAK qui rapporte cette définition, pressé d'arriver à la théorie des nombres complexes (agrégats avec différentes espèces d'éléments), ne précise pas davantage. En 1883, CANTOR ⁽¹⁾ la reprend — d'après « les leçons de WEIERSTRASS sur les fonctions analytiques » ⁽²⁾ — comme introduction à la sienne qu'il réexpose de façon plus systématique. « Le moment générateur qui lie l'agrégat avec le nombre défini par lui réside dans la sommation » (des éléments de l'agrégat). Cependant cette sommation n'est jamais que finie ; il y aurait la même « faute logique » à définir le nombre irrationnel comme somme infinie que comme limite de nombre rationnel puisque « la définition de la somme infinie n'est obtenue que par identification avec le nombre irrationnel auparavant défini » ⁽³⁾. L'inconvénient que le même nombre peut être représenté par une infinité d'agrégats s'évite aisément : il suffit de spécialiser le mode de constitution des agrégats, d'utiliser seulement, par exemple, comme éléments les fractions décimales ou les fractions continues. Mais il y a une autre difficulté : la sommation finie des éléments d'un agrégat (nécessaire, par exemple, pour la comparaison avec un autre nombre) doit

(1) CANTOR (XV), p. 190.

(2) *Ibid.*, p. 183.

(3) P. 184-185.

s'effectuer dans un ordre déterminé si l'agrégat est constitué par les termes d'une série non absolument convergente. Or la définition de WEIERSTRASS ne peut faire intervenir aucun ordre entre les éléments.

Aussi vaut-il mieux substituer à la notion d'agrégat celle de suite fondamentale introduite simultanément, en 1871, par CANTOR ⁽¹⁾ et HEINE. Avec toujours comme base l'ensemble des nombres rationnels, on considère l'ensemble des suites convergentes formées de ces nombres. Ce sera l'ensemble des nombres irrationnels d'ordre 1. On voit que tout nombre rationnel a y a un représentant : la suite $(a, a...)$. L'ordre et les quatre opérations sont aisés à définir : deux éléments sont égaux si la différence $a_n - a'_n$ des termes de même rang des suites fondamentales tend vers 0 avec $\frac{1}{n}$; un élément en précède un autre si $a_n - a'_n$, à partir d'un certain rang, reste plus grand qu'un nombre positif rationnel : un de ces deux cas se présente forcément pour tout couple d'éléments. La somme, la différence, le produit de deux éléments ont une suite fondamentale composée respectivement des sommes, différences ou produits des termes de même rang dans les suites. D'une façon générale toute relation algébrique $\Phi(\alpha, \beta, \gamma, ...)$ entre éléments se traduit en relations algébriques $F(a_n, b_n, c_n, ...)$ entre nombres rationnels. Si des éléments $\alpha_1, \alpha_2, ... \alpha_n, ...$ forment une suite convergente ⁽²⁾, il est toujours possible de former avec les nombres rationnels de leurs suites respectives une seule suite fondamentale qui définit un nouvel élément β appelé limite de la suite $\alpha_1, \alpha_2, ...$: on voit que $\beta - \alpha_n$ peut être, d'après les définitions données, rendu aussi petit qu'on veut pour n assez grand. Toutefois, si l'on considère l'aspect « conceptuel » des opérations, la

(1) CANTOR (V), p. 92-96 ; HEINE (II), p. 172.

(2) Une suite d'éléments α_i est convergente si la différence $\alpha_{m+n} - \alpha_m$ — qu'on a appris à définir — est plus petite qu'un nombre donné quelconque pour n assez grand.

suite convergente $\alpha_1, \alpha_2, \dots \alpha_n \dots$ engendre en réalité un élément d'un autre ordre : on superpose donc au domaine des nombres irrationnels d'ordre 1, un domaine d'ordre 2 formé des suites fondamentales d'éléments d'ordre 1 ou d'ordre 0 (nombres rationnels), et ainsi de suite indéfiniment, les domaines d'ordre ≥ 1 ayant la propriété de se recouvrir tous exactement (chaque élément de l'un d'eux a un représentant dans tous les autres). CANTOR voit à sa définition l'avantage qu'elle décrit exactement le procédé qui donne pratiquement les nombres irrationnels dont s'occupe l'analyse. Ainsi pour « décrire le contenu intellectuel de la proposition $\sin \frac{\pi}{2} = 1$ on peut se donner $\frac{\pi}{2}$ et ses puissances par les suites ⁽¹⁾ :

pour $\frac{\pi}{2}$: (α_v)

pour $\left(\frac{\pi}{2}\right)^{2m+1}$: (α_v^{2m+1}) où $\alpha_v = 2 \sum_{n=0}^v \frac{(-1)^n}{2n+1}$

Enfin : $\sin \frac{\pi}{2} : \left(\sum_{m=0}^p (-1)^m \frac{\left(\frac{\pi}{2}\right)^{2m+1}}{2m+1} \right)$

c'est-à-dire $\sin \frac{\pi}{2}$ est défini grâce à une suite fondamentale de 2^e ordre et par cette proposition est exprimée l'égalité entre le nombre rationnel 1 et un nombre $\sin \frac{\pi}{2}$ donné par une suite fondamentale de 2^e ordre. De façon analogue, le contenu intellectuel de formules compliquées comme celles de la théorie des fonctions θ se peut décrire avec précision et une relative simplicité ⁽²⁾. »

Mais cette « distinction conceptuelle » entre les différents

(1) CANTOR et HEINE représentent une suite fondamentale de terme général a_n par (a_n) .

(2) CANTOR (XV), p. 189.

ordres de nombres irrationnels avait, dans le mémoire de 1871, un but précis. Définissons comme points rationnels sur un axe les points dont l'abscisse est donnée par un nombre rationnel. Tout point de l'axe « s'il nous est connu par une construction quelconque » ⁽¹⁾, il est possible de donner une suite fondamentale (a_n) , telle que les points d'abscisse a_n se rapprochent indéfiniment de lui. Inversement, ce sera un axiome « parce qu'il est dans sa nature de ne pouvoir être démontré de façon générale » ⁽²⁾ — l'axiome de CANTOR — qu'à tout nombre irrationnel correspond un point de la droite ayant ce nombre pour abscisse. Si l'on considère maintenant l'ensemble A' des points limites d'un ensemble A (son 1^{er} dérivé, c'est-à-dire l'ensemble des points tels qu'il y ait, aussi près d'eux qu'on veut, toujours des points de cet ensemble) ⁽³⁾, puis l'ensemble A'' des points limites de A' (2^e dérivé) et ainsi de suite, l'ensemble A pourra être caractérisé par le nombre possible de ces opérations successives de *dérivation* : on devra s'arrêter à un ensemble $A^{(v)}$ s'il ne comporte qu'un nombre fini de points (donc pas de point limite). L'ensemble A sera dit de v^e espèce. « Un exemple d'ensemble de v^e espèce est déjà offert par un seul point, si son abscisse est donnée par un nombre irrationnel d'ordre v , satisfaisant à des conditions faciles à poser. Si l'on résoud, en effet, ce nombre en ses éléments, les termes d'ordre $v-1$ de la suite fondamentale correspondante, puis ceux-ci dans leurs éléments constitutifs d'ordre $v-2$, etc...., on obtient enfin une infinité de nombres rationnels, l'ensemble des points correspondant est un ensemble de v^e espèce ⁽⁴⁾. »

Ainsi c'est bien la théorie des ensembles naissante qui dicte à CANTOR cette transformation de la définition de

(1) CANTOR (V), p. 96.

(2) P. 97.

(3) Pour la définition des points limites et de l'opération de dérivation, voir ci-après, p. 56 sq.

(4) P. 98.

WEIERSTRASS. La notion de suite fondamentale a comme intérêt direct de donner le corrélat analytique de l'opération de dérivation. Celle-ci elle-même se trouvait exigée à son tour par les recherches faites à la suite de RIEMANN et de HEINE sur l'unicité de la représentation d'une fonction par une série trigonométrique.

* * *

On sait l'importance centrale des travaux sur les séries trigonométriques dans le développement de l'analyse au XIX^e siècle : c'est grâce à eux que la notion générale de fonction a pu se dégager, c'est à leur occasion que les principales extensions de l'intégration ont été effectuées de LEJEUNE-DIRICHLET à RIEMANN et même à LEBESGUE ⁽¹⁾. D'où la mise en jeu de la théorie des ensembles, à propos tant du but des recherches — théorie générale des fonctions, — que des moyens utilisés — intégrales de FOURIER.

HANKEL fait dater les mathématiques modernes « de l'instant où DESCARTES, à partir du traitement purement algébrique des équations, s'avança jusqu'à étudier les variations de grandeur que subit une expression algébrique lorsqu'une grandeur désignée en elle de façon générale, parcourt une suite continue de valeurs. » ⁽²⁾ C'est cependant LEIBNIZ qui introduisit le mot même de fonction, et s'il semble en avoir eu une notion assez générale, comme celle de dépendance entre deux grandeurs géométriques données, — « *parametri seu rectæ magnitudine constantes... censentur... indifferentiabiles, quemadmodum et ipsa recta tangens vel aliæ nonnullæ functiones ab ea pendentes*, v. g.

(1) Dès la préface de la 1^{re} édit. de ses *Leçons sur l'intégration* (1907), LEBESGUE signale comme une « très importante application » de sa définition « l'étude du développement trigonométrique des fonctions ». En effet, « pour qu'une fonction admette une série de FOURIER, il faut et il suffit qu'elle soit sommable ». LEBESGUE (II), p. 31.

(2) HANKEL (I), p. 45.

perpendiculares ad tangentiam ab axe ad curvam ductæ» ⁽¹⁾, — du moins les grandeurs elles-mêmes, puisqu'il s'agit de courbes connues, étaient-elles définies par des expressions analytiques simples, de sorte que JOHANN BERNOULLI pouvait se réclamer de lui pour sa définition, reprise ensuite textuellement par EULER et classique pour le XVIII^e siècle : « on appelle fonction d'une grandeur variable, une quantité composée de quelque manière que ce soit de cette grandeur variable et de constantes » ⁽²⁾. Ici c'est la composition qui définit la fonction, non la dépendance posée entre termes donnés extra-analytiquement. D'où la limitation des fonctions aux expressions analytiques connues : mais le décalage avec la première définition exigeait un élargissement. Déjà EULER distinguait ⁽³⁾ entre *curvae continuæ*, celles dont le domaine coïncide avec le domaine des expressions analytiques, et *curvae discontinuæ, seu mixtæ seu irregulares*, les « courbes mécaniques » de DESCARTES données graphiquement, par exemple formées de morceaux bout à bout de courbes continues. Le procédé même semblait exclure une définition analytique quelconque (non plus seulement donc limitée aux expressions connues) puisque, comme le note LEBESGUE, « on admettait que deux expressions analytiques égales dans un intervalle sont égales partout » ⁽⁴⁾. Mais, à ce renoncement devant le mécanique, le mouvement spontané des mathématiques, en tant qu'expression — et solution — de tout problème posé précisément obligeait de passer outre.

En 1753, Daniel BERNOULLI et EULER étudient le mouvement des cordes vibrantes. Il s'agit de donner à l'équation aux dérivées partielles

(1) *Acta Erudit. Lips.*, 1692.

(2) Joh. BERNOULLI, *Opera omnia*, II, p. 241 (cité par HANKEL, *ibid.*).

(3) *Opusc. math.*, t. I, p. 41, cité par LEBESGUE (II), p. 21 n.

(4) LEBESGUE, *ibid.*

$$\frac{\partial^2 y}{\partial t^2} = K^2 \frac{\partial^2 y}{\partial x^2},$$

la solution la plus générale satisfaisant à un certain nombre de conditions aux limites. En particulier, pour $t = 0$, la fonction $y = f(x, t)$ doit représenter la position initiale de la corde, or, remarque d'ALEMBERT, « la manière ordinaire, pour ne pas dire l'unique, de faire sortir une corde de son état de repos, c'est de la prendre par un de ses points et de la tendre en la tirant, ce qui lui donne la figure de deux lignes droites qui font un angle entre elles » (1). Il est donc « contre les règles de l'analyse » de prétendre, comme faisait EULER, qu'une série trigonométrique pourrait fournir la solution. Cependant en 1807 FOURIER donnait un procédé régulier pour construire cette série ; si l'on pose :

$$f(x) = \sum_{k=0}^{\infty} (a_k \cos kx + b_k \sin kx)$$

en multipliant les deux membres par $\sin kx$ ou $\cos kx$ et en intégrant terme à terme de $-\pi$ à $+\pi$, on obtient :

$$b_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\alpha) d\alpha \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(\alpha) \cos k\alpha d\alpha$$

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(\alpha) \sin k\alpha d\alpha.$$

L'originalité de FOURIER, remarque SACHSE (2), n'est pas — contrairement à ce que dit RIEMANN — dans le procédé de calcul des coefficients, déjà utilisé par LAGRANGE et EULER, aux travaux desquels d'ailleurs FOURIER renvoie lui-même, mais d'avoir vu « qu'une série trigonométrique, avec des coefficients ainsi déterminés, peut représenter une fonction arbitraire quelconque » : la méthode semblait tellement sûre pour résoudre le problème que pendant un demi-siècle

(1) LEBESGUE, *ibid.*

(2) SACHSE (1), p. 47.

séries trigonométriques et séries de FOURIER furent identifiées, jusqu'à ce que HEINE, en 1870, s'avisât le premier de faire la distinction. D'où la mise au premier plan, pendant presque toute cette période, des études sur l'intégration : non seulement le calcul des coefficients mais la convergence de la série s'y trouvent subordonnés pour DIRICHLET. Quelles conditions, remplies par la fonction arbitraire, suffisent pour qu'une certaine intégration soit possible, tel est le problème dont lui et ses disciples s'occupent. Que résulte-t-il pour la marche d'une fonction de cette propriété d'être représentable par une série trigonométrique quelconque — et cette représentation est-elle univoque, c'est la seconde question posée, de RIEMANN à CANTOR, et qui, attaquant de front la notion de fonction, devait mener directement à la théorie des ensembles.

FOURIER lui-même ne s'est pas préoccupé de démontrer la convergence de sa série : il lui suffit que l'on puisse voir dans des cas particuliers que la série converge effectivement vers la valeur de la fonction correspondante. La première démonstration due à CAUCHY repose sur une confusion entre convergence et absolue convergence : la série de FOURIER converge parce que le rapport entre son terme général et $\frac{A \sin n x}{n}$ (A étant une constante) tend vers $+1$ lorsque n croît indéfiniment. « Mais, objecte DIRICHLET en 1829 ⁽¹⁾, il est facile de donner l'exemple de séries, dont les termes ne sont pas tous positifs, telles que l'une soit convergente, l'autre divergente quoique le rapport des termes correspondants tende vers $+1$:

$$\sum_n \frac{(-1)^n}{\sqrt{n}} \quad \text{et} \quad \sum_n \frac{(-1)^n}{\sqrt{n}} \left[1 + \frac{(-1)^n}{\sqrt{n}} \right]. »$$

La méthode de DIRICHLET consiste à mettre la somme des n premiers termes de la série de FOURIER sous la forme :

(1) LEJEUNE-DIRICHLET (I), p. 158.

$$\frac{1}{\pi} \int_{-\pi}^{+\pi} \varphi(\alpha) \frac{\sin \frac{2n+1}{2}(\alpha-x)}{2 \sin \frac{\alpha-x}{2}} d\alpha \quad (1)$$

($\varphi(\alpha)$ étant la fonction arbitraire). « Tout se réduit à déterminer la limite dont cette intégrale approche quand n croît indéfiniment ». Lorsque $\varphi(\alpha)$ est continue, monotone et finie, la limite est bien $\varphi(x)$. Mais la notion d'intégrale telle que CAUCHY l'avait précisée permet d'aller plus loin :

1° l'intégrale d'une fonction dans un intervalle étant la somme des intégrales étendues à un nombre fini d'intervalles partiels en lesquels le premier est divisé, on peut accorder à $\varphi(x)$ un nombre fini de maxima et de minima ;

2° Si $\int_a^x f(\alpha) d\alpha$ tend vers une limite déterminée lorsque x tend vers b , point singulier pour f (point de discontinuité ou point où f est infinie), CAUCHY définit par cette limite l'intégrale $\int_a^b f(\alpha) d\alpha$. En particulier si $f(x)$ tend vers une limite déterminée, $f(b-0)$ ou $f(b+0)$ (notation de DIRICHLET), pour x tendant vers b en croissant ou décroissant, c'est-à-dire si b est un point de discontinuité de 1^{re} espèce (condition toujours réalisée, ainsi que le remarque RIEMANN lorsque $f(x)$ n'a qu'un nombre fini de maxima et de minima), les deux intégrales $\int_a^{b-0} f(\alpha) d\alpha$ et $\int_{b+0}^c f(\alpha) d\alpha$ auront des valeurs bien déterminées, l'intégrale $\int_a^c f(\alpha) d\alpha$ (avec $a < b < c$) sera définie comme leur demi-somme.

En utilisant donc seulement les résultats déjà connus, DIRICHLET peut affirmer que la série (1) est convergente si $\varphi(x)$ remplit les trois conditions :

1° rester finie dans l'intervalle $-n \dots +n$;

2° n'avoir qu'un nombre fini de discontinuités (de 1^{re} espèce) ;

3° n'avoir qu'un nombre fini de maxima et de minima.

Pour que la série représente effectivement $\varphi(x)$, RIEMANN remarque (1) qu'il est nécessaire d'ajouter une 4^e condition : $\varphi(x)$ est égale à la valeur moyenne de ses deux valeurs limites $\varphi(x + 0)$ et $\varphi(x - 0)$ aux points de discontinuité ; sans quoi il faut renoncer à la coïncidence entre la série et la fonction en ces points (2). En revanche, aucune des trois premières conditions n'est irréductible : « il nous resterait à considérer les cas où les suppositions, que nous avons faites sur le nombre des solutions de continuité et sur celui des valeurs maxima et minima, cessent d'avoir lieu. Ces cas singuliers peuvent être ramenés à ceux que nous venons de considérer... Mais la chose, pour être faite avec toute la clarté qu'on peut désirer, exige quelques détails liés aux principes fondamentaux de l'analyse infinitésimale et qui seront exposés dans une autre note (3). »

Cette promesse de révision des principes, c'est-à-dire d'une nouvelle théorie de l'intégration, n'a pas été tenue. La seule restriction levée fut en 1837 (4) celle dont justement il n'était pas question en 1829, l'obligation de rester finie pour $\varphi(x)$. Si $\int_x^y \varphi(\alpha) d\alpha$ est une fonction finie et continue des variables x et y , ($\varphi(\alpha)$ pouvant devenir infinie entre x et y), quelles que soient les valeurs prises par x et y entre $-\pi$ et π (c'est-à-dire si $\varphi(\alpha)$ est intégrable au sens CAUCHY entre $-\pi$ et $+\pi$), l'intégrale (1) prise non entre $-\pi$ et π mais entre des valeurs arbitrairement proches encadrant la valeur rendant $\varphi(x)$ infinie, peut être rendue aussi petite qu'on le veut. Il suffit alors, suivant la remarque de LIPSCHITZ, s'il n'y a qu'un nombre fini de points où $\varphi(x)$ est infinie, de les entourer d'intervalles arbitrairement petits : leur contribution

(1) RIEMANN (I). ..

(2) Plus tard, SCHÄFFLI (I) et DU BOIS REYMOND (V) devaient revenir sur le caractère indépendant (par rapport aux autres) de cette nouvelle condition.

(3) LEJEUNE-DIRICHLET (I), p. 169.

(4) LEJEUNE-DIRICHLET (III).

dans l'intégrale (1) prise de $-\pi$ à π est négligeable. La condition 1^o se transforme donc en : n'être infinie qu'en un nombre fini de points. Pour l'affaiblissement de la condition 2^o DIRICHLET donne une indication dès 1829 : « il est nécessaire que la fonction $\varphi(x)$ soit telle que, si l'on désigne par a et b deux quantités quelconques comprises entre $-\pi$ et π , on puisse toujours placer entre a et b d'autres quantités r et s assez rapprochées pour que la fonction reste continue dans l'intervalle de r à s . On sentira facilement la nécessité de cette restriction en considérant que les différents termes de la série sont des intégrales définies et en remontant à la notion fondamentale des intégrales. On verra alors que l'intégrale d'une fonction ne signifie quelque chose qu'autant que la fonction satisfait à la condition précédemment énoncée. On aurait un exemple d'une fonction qui ne remplit pas cette condition, si l'on supposait $\varphi(x)$ égale à une constante déterminée c lorsque la variable x obtient une valeur rationnelle, et égale à une autre constante d lorsque cette variable est irrationnelle » (1). L'exemple ne suffit pas à éclaircir l'obscurité du passage : s'il est évident qu'une fonction dont les discontinuités sont partout denses dans un intervalle n'est pas intégrable (au sens CAUCHY ou RIEMANN) dans cet intervalle, la densité (la condition imposée revient à dire que l'ensemble des points de discontinuité est non dense entre $-\pi$ et π) ne joue cependant pas de rôle dans l'intégration telle qu'on l'envisageait alors. Le but était d'enfermer les points de discontinuité dans des intervalles dont la contribution puisse être négligée dans l'intégrale totale. Il semble qu'il y ait eu dans l'esprit de DIRICHLET interférence entre deux considérations — que la théorie des ensembles devait distinguer : — celle de la mesure, qui devait mener à la notion de groupe intégrable, introduite par DU BOIS REYMOND ; celle de la répartition des points de

(1) LEJEUNE-DIRICHLET (I), p. 169.

discontinuité autour de points limites, notion d'ensemble réductible précisée par CANTOR. Mais pour utiliser la 1^{re}, l'intégrale de RIEMANN était indispensable. Aussi dans son commentaire de ce passage LIPSCHITZ choisit-il la seconde, sans même apercevoir, ainsi que le note MONTEL ⁽¹⁾, qu'il transforme la condition de DIRICHLET. L'explication n'est d'ailleurs guère précise : en supposant que l'ensemble des points de discontinuité ne comporte qu'un nombre fini de points limites, on peut enfermer ceux-ci dans un nombre fini d'intervalles de longueur arbitrairement petite ; dans les intervalles qui subsistent, la fonction n'a qu'un nombre fini de discontinuités ; pour les premiers intervalles les intégrales correspondantes peuvent être rendues aussi petites qu'on le veut ; en effet « chacune de ces intégrales est la somme d'une infinité d'intégrales constituées de manière que cette somme même puisse être rendue aussi petite qu'on le veut » ⁽²⁾. On doit voir là, sans doute comme le propose LEBESGUE ⁽³⁾, l'application d'une définition : l'intégrale d'une fonction $f(x)$ entre deux points limites α, β de ses points de discontinuité est :

$$\int_{\alpha}^{\beta} f(x) dx = \lim_{\substack{\varepsilon = 0, \\ \eta = 0}} \int_{\alpha + \varepsilon}^{\beta - \eta} f(x) dx$$

lorsque cette limite existe, et il n'y a d'intégrale que si cette limite existe.

Contrairement à l'opinion exprimée par RIEMANN ⁽⁴⁾, LIPSCHITZ montre également ⁽⁵⁾ que la 3^e condition peut être supprimée si en tous les points où la fonction oscille, l'oscillation $\omega = |f(x + \delta) - f(x)|$ est telle que :

$$\omega < k \delta \alpha$$

(1) LIPSCHITZ (I), p. 284, n. 2.

(2) *Ibid.*, p. 284.

(3) LEBESGUE (I), p. 10.

(4) RIEMANN (I), p. 224 : « Si la dernière supposition (que la fonction n'a pas une infinité de maxima et de minima) est abandonnée, les deux théorèmes de DIRICHLET sont insuffisants pour décider de la question. »

(5) LIPSCHITZ (I), p. 285-295.

(k étant une constante, α un nombre > 0), exige encore trop forte puisque, comme le fait remarquer SACHSE ⁽¹⁾, LIPSCHITZ se sert seulement de celle-ci :

$$\lim_{\delta=0} \omega \log \delta = 0$$

— « hac in re sunt nervi demonstrationis nostræ ⁽²⁾ » — adoptée ensuite par U. DINI. Ici encore interviennent des notions de la théorie des ensembles : « ou bien on peut intercaler entre a et b des nombres $r - \delta$ et $r + \delta$ tels qu'il y ait un nombre fini de maxima et minima dans $(a, r - \delta)$ et $(r + \delta, b)$, mais un nombre infini entre $r - \delta$ et $r + \delta$, quelque petite que soit la distance 2δ de ces deux nombres ; ou bien quelles que soient les quantités r et s situées dans l'intervalle (a, b) et à une distance finie l'une de l'autre, on ne peut jamais trouver un nombre fini de maxima et de minima dans l'intervalle (r, s) ; ou bien l'intervalle considéré contient un nombre fini d'intervalles non nuls dans lesquels l'une ou l'autre hypothèse est réalisée ⁽³⁾. » Répartition en trois cas, qui, comme le note MONTEL ⁽⁴⁾, ne joue d'ailleurs aucun rôle dans la démonstration. Mais elle éclaire la confusion précédente entre ensemble à 1^{er} dérivé fini et ensemble non dense. Si, en effet, on remplace la définition du premier cas par celle de DIRICHLET : ensemble des maxima ou minima non dense, et dans celle du troisième cas le mot fini par « fini ou infini », la répartition est cette fois exhaustive. C'est la considération précise du rapport entre points et intervalles les enfermant qui fait défaut à DIRICHLET et LIPSCHITZ : elle ne devait s'éclaircir, grâce à la notion d'ensemble parfait non dense, par exemple, qu'après les travaux de RIEMANN sur l'intégration.

(1) SACHSE (I), p. 62.

(2) LIPSCHITZ (I), p. 294, texte original de LIPSCHITZ ; les p. renvoient à la traduction MONTEL.

(3) P. 286.

(4) *Ibid.*, p. 286, n. 1.

La partie positive du mémoire de RIEMANN commence, en effet, par une nouvelle définition de l'intégrale, dont les propriétés permettent de démontrer à la fois un des lemmes fondamentaux, et directement, quelques-uns des principaux résultats, comme le théorème que lorsque la fonction est intégrable dans ce nouveau sens, les coefficients de la série trigonométrique correspondante deviennent infiniment petits. RIEMANN transforme, ainsi que le note LEBESGUE ⁽¹⁾, le procédé opératoire donné par CAUCHY pour calculer l'intégrale d'une fonction continue en définition de l'intégrale d'une fonction quelconque. Si la somme :

$$S = \delta_1 f(x_1) + \delta_2 f(x_2) + \dots + \delta_n f(x_n)$$

(les δ_i représentent des intervalles partiels en lesquels est divisé l'intervalle d'intégration $(a \ b)$, les x_i des points quelconques à l'intérieur des δ_i) tend vers une limite, quels que soient les intervalles δ_i « lorsqu'ils deviennent tous infiniment petits, sa valeur s'appelle :

$$\int_a^b f(x) \, dx.$$

Si... [la somme S] ... n'a pas cette propriété $\int_a^b f(x) \, dx$ n'a aucune signification ⁽²⁾. » D'où une conséquence qui montrera l'élargissement par rapport à CAUCHY : si on appelle ω_i l'oscillation de $f(x)$ dans l'intervalle δ_i (c'est-à-dire la différence entre la limite supérieure et la limite inférieure dans cet intervalle), la condition nécessaire et suffisante pour que S ait une limite, c'est-à-dire pour que $f(x)$ soit intégrable, lorsque $f(x)$ reste finie dans $(a \ b)$ est que l'on puisse trouver une division de $(a \ b)$ en intervalles δ telle que les δ_i pour lesquels l'oscillation ω_i correspondante est supérieure à un nombre quelconque donné σ aient une longueur totale aussi petite qu'on veut. Ainsi se trouve préci-

(1) LEBESGUE (I), p. 24.

(2) RIEMANN (I), p. 225.

sée et étendue l'idée de DIRICHLET-LIPSCHITZ : non seulement il suffit que les points de discontinuité d'une fonction intégrable puissent être enfermés dans un nombre fini d'intervalles de longueur totale arbitrairement petite, c'est-à-dire forment ce que DU BOIS REYMOND devait appeler un groupe intégrable (ensemble évidemment non dense, mais en plus de mesure nulle, ce que n'avait pas distingué DIRICHLET, en tous cas plus général que l'ensemble, à 1^{er} dérivé fini, auquel s'était borné LIPSCHITZ), mais encore il n'est nécessaire que de considérer l'importance des discontinuités relativement aux intervalles où elles apparaissent : si les points où les sauts ⁽¹⁾ sont inférieurs à une quantité donnée forment un groupe intégrable, l'intégrale existe. Ainsi la série :

$$f(x) = \sum_{n=1}^{\infty} \frac{(nx)}{n^2},$$

où (nx) représente l'excès de nx sur l'entier le plus voisin (on pose $(nx) = 0$ lorsque nx est à égale distance de deux entiers), convergente pour tout x , est discontinue pour toute valeur rationnelle de x qui, sous forme de fraction irréductible, a un dénominateur pair ; les points de discontinuité sont partout denses. Mais, d'une part, les discontinuités sont de première espèce, d'autre part le nombre des sauts supérieurs à une grandeur donnée est fini. En réalité, RIEMANN n'a pas directement besoin du groupe intégrable : tout ce qui concerne son intégrale (lorsque la fonction à intégrer est finie) se traite, comme le montre son travail, à l'aide de propriétés d'intervalles. Si la théorie des ensembles est ici latente, elle n'a pas trouvé encore obligation d'apparaître.

(1) Le saut de $f(x)$ au point x est la limite vers laquelle tend

$$f(x + \varepsilon) - f(x - \eta)$$

lorsque ε et η tendent vers 0 : $f(x + 0) - f(x - 0)$.

A l'étude des conditions suffisantes — à laquelle s'étaient exclusivement consacrés ses prédécesseurs, — RIEMANN se propose de substituer celle des conditions nécessaires pour la représentabilité d'une fonction par une série trigonométrique, « et parmi celles-ci ensuite de choisir les conditions suffisantes ». SACHSE observe que le problème traité en fait est différent (et ceci expliquerait l'incorrection de la phrase puisque les conditions suffisantes comprennent évidemment toutes les conditions nécessaires) : il s'agit des conditions nécessaires pour qu'une fonction soit égale à une série trigonométrique partout où celle-ci converge. Mais peut-on parler encore de représentation, s'il y a, par exemple, dans un intervalle une infinité de points où la série ne converge pas ? SACHSE propose de se restreindre à un nombre fini d'exceptions. Le résultat fondamental obtenu par RIEMANN est que les conditions qu'il trouve sont également suffisantes, et « que, lorsque les coefficients de la série deviennent infiniment petits $\left[\text{avec } \frac{1}{n} \right]$, sa convergence pour une certaine valeur de x ne dépend que du comportement de la fonction au voisinage immédiat de cette valeur ». D'où la nécessité, d'une part, d'analyser ce voisinage — de faire donc intervenir implicitement la notion de point limite, — d'autre part, pour la représentation d'une fonction donnée, de rechercher des conditions suffisantes. RIEMANN montre que si les trois conditions de DIRICHLET transformées (1. la fonction ne doit pas être infinie ; 2. elle doit être intégrable au sens RIEMANN ; 3. elle doit ne pas comporter une infinité de maxima et de minima) sont remplies, les conditions nécessaires le sont également, les coefficients de la série sont ceux de FOURIER et la série converge en tous les points qui n'ont pas dans leur voisinage une infinité de points de discontinuité de la fonction. Reste l'hypothèse que les coefficients ne deviennent pas infiniment petits avec $\frac{1}{n}$ pour toute va-

leur de x : on a vu qu'elle ne peut pas se présenter lorsque $f(x)$ est intégrable. Dans les autres cas, et si $f(x)$ ne comporte pas une infinité de maxima et de minima, on peut appliquer, par un artifice immédiat, les raisonnements précédents : la série ne peut converger qu'en un nombre fini de points ; on ne peut guère parler de représentation.

Mais est-ce la propriété d'être égale, là où elle converge, à une fonction ne possédant qu'un nombre fini de maxima et de minima, qui oblige une série trigonométrique dont les coefficients ne tendent pas vers 0, à ne converger qu'en un nombre fini de points ? Plus généralement, que résulte-t-il de cette propriété pour l'aspect même de la série ? D'autre part, y a-t-il une différence de portée entre les raisonnements de RIEMANN pour l'établissement des conditions nécessaires — qui concernent un développement trigonométrique quelconque — et les démonstrations des conditions suffisantes, qui font intervenir les propriétés des coefficients de FOURIER ? Double façon d'arriver au problème de l'unicité du développement. Le procédé même de FOURIER garantit la réponse, sous réserve qu'on puisse intégrer terme à terme, ce qu'admettaient FOURIER et CAUCHY. Avec la découverte de la condition de convergence uniforme un nouvel examen s'impose.

En 1822, CAUCHY énonçait dans son *Cours d'Analyse* que « lorsque les différents termes d'une série sont des fonctions d'une même variable x , continues par rapport à cette variable dans le voisinage d'une valeur particulière pour laquelle la série est convergente, la somme S de la série est aussi, dans le voisinage de cette valeur particulière, fonction continue de x » (1). Théorème dont, en 1826, ABEL montrait « qu'il souffre des exceptions » (2), par exemple la série :

(1) CAUCHY (I), p. 132.

(2) ABEL (I), p. 312.

$$f(\varphi) = \sin \varphi - \frac{\sin 2\varphi}{2} + \frac{\sin 3\varphi}{3} \dots$$

discontinue pour toute valeur $\varphi = (2m + 1)\pi$ (m étant un nombre entier). En 1848, SEIDEL remarquait qu'il était également « en contradiction avec les travaux de DIRICHLET puisque la série trigonométrique converge même quand elle représente une fonction discontinue » (1). D'une façon générale, « un agrégat de grandeurs continues ne pourrait jamais représenter des grandeurs discontinues si on fait intervenir la forme de l'infini ». En réalité, « si on a une série convergente qui représente une fonction discontinue de x et dont les termes sont des fonctions continues de x , on doit, dans le voisinage de la valeur où la fonction *saute*, trouver des valeurs de x pour lesquelles la série converge *arbitrairement lentement* ». Si la vitesse de convergence est, au contraire, égale pour tous les points d'un intervalle, c'est-à-dire, si quel que soit x , on a :

$$\text{pour } n > N \quad | u_{n+m}(x) - u_n(x) | < \varepsilon$$

(ε étant un nombre arbitraire donné auquel on fait correspondre l'entier N , m étant quelconque), il y a convergence uniforme (2) : WEIERSTRASS démontrait peu après que c'était la condition nécessaire et suffisante pour qu'on puisse intégrer terme à terme. En complément du théorème de SEIDEL, DU BOIS REYMOND (3) établissait que, même aux points de continuité d'une fonction, la convergence de sa série peut n'être pas uniforme : de sorte que, même pour les fonctions continues, l'unicité du développement trigonométrique n'est pas garantie par le procédé de FOURIER.

(1) SEIDEL (I), p. 382.

(2) CAUCHY rectifia lui-même son erreur dans une note aux *Comptes rendus*, en 1853. « Il est facile de voir comment on doit modifier l'énoncé du théorème pour qu'il n'y ait plus lieu à aucune exception » et il donne exactement la définition de la convergence uniforme (CAUCHY (II), p. 455).

(3) DU BOIS REYMOND (X).

HEINE (1) introduit la notion de convergence uniforme « en général », c'est-à-dire à l'exception des voisinages immédiats d'un nombre fini de points de discontinuité. Une fonction finie, ne possédant pas une infinité de maxima et de minima et continue « en général » (c'est-à-dire sauf en un nombre fini de points) ne peut être développée que d'une façon en série trigonométrique qui soit, en général, uniformément convergente. Autrement dit, une série trigonométrique en général uniformément convergente ne peut représenter que si tous ses coefficients sont nuls. HEINE le démontre en intégrant deux fois terme à terme dans les intervalles $(x_n \dots x_{n+1})$, à l'intérieur desquels il y a convergence uniforme, tandis qu'elle cesse pour les extrémités : on obtient des fonctions linéaires $F(x) = a_n x_n + b_n$.

Or, en vertu d'une des conditions nécessaires établies par RIEMANN, on doit avoir :

$$\lim_{\alpha \rightarrow 0} \frac{F(x+\alpha) + F(x-\alpha) - 2F(x)}{2\alpha} = 0 \quad (2)$$

Il en résulte que tous les a_n et les b_n sont égaux entre eux, et par suite tous nuls (puisque pour $x = 0$ on a $b_0 = 0$). La série qui représente $F(x)$ est donc nulle partout : comme elle est également partout uniformément convergente, on peut lui appliquer le procédé de FOURIER, tous ses coefficients sont nuls.

L'hypothèse de la convergence uniforme n'est pas essentielle au raisonnement si on utilise un autre résultat de RIEMANN, établi avec la seule supposition que les coefficients tendent vers 0 avec $\frac{1}{n}$.

CANTOR (2) commence par démontrer que c'est le cas pour toute série trigonométrique convergente dans tout un in-

(1) HEINE (II), p. 353.

(2) CANTOR (I) et (III).

tervalle (à l'exception de points isolés) répondant ainsi à la première question laissée en suspens par RIEMANN. Mais le théorème n'est même pas indispensable comme fit observer KRONECKER ⁽¹⁾ en reprenant un artifice que RIEMANN avait d'ailleurs déjà utilisé. Une des conditions nécessaires de RIEMANN est que, si on pose :

$$F(x) = b_0 \frac{x^2}{4} - \sum_{n=1}^{\infty} \frac{a_n \sin nx + b_n \cos nx}{n^2} \quad (3)$$

(série obtenue par deux intégrations consécutives terme à terme de la série donnée $f(x)$) on a pour la dérivée seconde généralisée de $F(x)$:

$$\lim_{\alpha \rightarrow 0} \frac{F(x+2\alpha) + F(x-2\alpha) - 2F(x)}{4\alpha^2} = f(x) \quad (4)$$

Ici la limite doit être 0. Or SCHWARZ ⁽²⁾ démontre que lorsque la dérivée seconde généralisée d'une fonction est nulle, la fonction est linéaire : on retrouve donc les conditions où se plaçait HEINE. S'il n'y a pas de points d'exception (points de discontinuité — c'est-à-dire ici points où $f(x)$ cesse d'être = 0 — ou points de non convergence pour la série), on voit sans peine que l'équation (3) devient :

$$K = \sum_{n=1}^{\infty} \frac{a_n \sin nx + b_n \cos nx}{n^2}$$

le second membre est évidemment uniformément convergent lorsque les coefficients a_n et b_n tendent vers 0 avec $\frac{1}{n}$, on peut donc appliquer le procédé FOURIER, et l'on trouve $a_n = b_n = 0$ quel que soit n . S'il y a des exceptions, le raisonnement de HEINE permet à CANTOR, grâce à la relation (2), de jeter en quelque sorte un pont entre les intérieurs de deux intervalles séparés par l'une de ces exceptions : les deux

(1) CANTOR (III), p. 84.

(2) SCHWARZ (I).

fonctions linéaires correspondantes sont identiques. Donc un intervalle ne comportant à son intérieur qu'un nombre fini de points exceptionnels correspondra à une seule fonction linéaire : ce seront alors les points limites des points exceptionnels qui sépareront deux intervalles semblables, et qui seront surmontés de la même façon grâce à la relation (2) : et ainsi de suite. « Une fonction discontinue $f(x)$ qui pour toutes les valeurs de x , qui correspondent aux points d'un ensemble de n^{e} espèce donné dans l'intervalle $(0 \dots 2n)$, est différente de 0 ou indéterminée, pour toutes les autres valeurs de x , est nulle, ne peut pas être représentée par une série trigonométrique ⁽¹⁾. »

Tel est l'acte de naissance, au moins de la théorie cantorienne des ensembles. On a vu comment s'y inscrit la définition du continu numérique. Mais l'important n'est pas là : la notion d'ensemble, la distinction entre les modes de répartition des points dans un intervalle avaient été imposées antérieurement par la théorie des fonctions. Aux environs de 1870 les travaux se multiplient qui s'en servent : HANKEL, à propos des *Fonctions infiniment souvent oscillantes et discontinues*, énonce son principe de condensation des singularités et définit oscillation en un point et oscillation dans un intervalle. DU BOIS REYMOND prétend avoir, dès cette époque, utilisé la distinction entre *panthies* (ensembles partout denses) et *apanthies* (ensembles non denses). Il n'y a théorie des ensembles qu'à l'apparition, non de notions, mais d'un mode de raisonnement original : tel est celui dont CANTOR faisait pour la première fois usage en généralisant le théorème de HEINE ; de lui — que l'on retrouvera sous le nom d'itération transfinie — dépend une partie essentielle de l'édifice futur.

(1) CANTOR (V), p. 101.

* * *

Un autre thème intervient cependant avec les travaux de DU BOIS REYMOND sur la croissance des fonctions. A propos de l'intégrale de FOURIER :

$$\int_{-\pi}^{\pi} f(x) \cos n(x-a) dx$$

dans le cas où n est une grandeur continue, il est amené à comparer les limites vers lesquelles tendent des expressions lorsque les variables y deviennent infinies de façon indépendante. Pour $z = f(xy)$, par exemple, il faudra distinguer les deux limites ⁽¹⁾ :

$$z_1 = \lim_{x \infty} \lim_{y \infty} f(xy) \quad \text{et} \quad z_2 = \lim_{y \infty} \lim_{x \infty} f(xy)$$

De même pour les façons de tendre vers 0. Après un essai de traitement géométrique, il jette les bases d'un *calcul infinitaire* permettant de comparer la croissance de deux fonctions différentes. L'ordre entre elles sera établi par les relations :

$$\begin{aligned} f(x) \prec \varphi(x) & \quad \text{si} \quad \lim_{x \infty} \frac{f(x)}{\varphi(x)} = 0 \\ f(x) \sim \varphi(x) & \quad \text{si} \quad \lim_{x \infty} \frac{f(x)}{\varphi(x)} = \text{une quantité finie} \end{aligned}$$

Ainsi pourrait se constituer un continu infinitaire, premier exemple d'ensemble, étudié mathématiquement, dont les éléments ne soient ni des points ni des nombres ⁽²⁾. « Nous nous proposons de développer quelques vues générales... concernant la suite continue des fonctions ordonnées sui-

(1) DU BOIS REYMOND (I).

(2) A peu près à la même époque THOMAE étudiait les ordres de décroissance de fonctions tendant vers 0.

vant les limites de leurs quotients et l'analogie de cette suite avec la suite des nombres réels ⁽¹⁾. » Aux nombres positifs correspondent les fonctions devenant infinies, aux nombres négatifs les fonctions devenant nulles, à 0 les fonctions restant finies. Il est possible de jalonner ce continu, comme au moyen d'entiers, avec ce que BOREL ⁽²⁾ devait appeler une « échelle de types croissants ». Il suffit de prendre une fonction croissante, l'exponentielle, par exemple, et son inverse la fonction logarithmique, et d'itérer indéfiniment l'une et l'autre :

$$l x = \underbrace{\log \log \dots \log x}_{\gamma} \quad \text{et} \quad e^{e^{\dots e^x}}$$

Entre deux ordres infinitaires différents on peut toujours insérer une infinité d'ordres intermédiaires, il suffit de multiplier la fonction d'ordre inférieur par une suite de fonctions croissant encore moins vite ; par exemple x^μ multiplié par $l x$, $l x.l_2 x$,... donnera des ordres tous inférieurs à celui de $x^{\mu+1}$. Mais ici s'arrête l'analogie avec le continu numérique. Le mode de raisonnement original donnant ici un début de théorie des ensembles, le procédé de la diagonale, montre en effet, que d'abord l'axiome d'ARCHIMÈDE n'est pas satisfait : étant donnée une suite infinie quelconque d'ordres infinitaires, il est toujours possible de construire une fonction d'ordre inférieur à tous les éléments de la suite. Si l'on prend une suite de courbes croissantes, en les coupant par des parallèles à oy, on voit que la courbe joignant l'intersection de la 1^{re} courbe et de la 1^{re} parallèle avec celle de la 2^e courbe et de la 2^e parallèle, etc... répond à la question (fig. 1). D'autre part ⁽³⁾, étant données deux suites

(1) DU BOIS REYMOND (II), p. 338.

(2) BOREL (I), p. 114.

(3) DU BOIS REYMOND (VII).

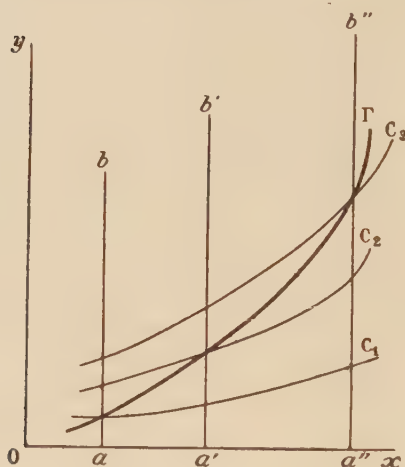


FIG. I.

de fonctions balayant le même intervalle infinitaire, il est impossible de représenter la même valeur de cet intervalle dans les deux suites. Soit, par exemple :

$$x^a \text{ et } (x \log x)^b \left\{ \begin{array}{l} \text{pour } a \text{ et } b < \frac{1}{10} \text{ les deux fonctions sont } < x^{\frac{1}{4}} \\ \text{pour } a \text{ et } b > \frac{9}{10} \text{ toutes deux } > x^{\frac{3}{4}} \end{array} \right.$$

cependant quels que soient a et b on ne peut jamais avoir :

$$x^a \sim (x \log x)^b$$

D'où l'affirmation, assez confuse, par DU BOIS REYMOND ⁽¹⁾ que le continu fonctionnel est plus riche, de puissance plus élevée que le continu numérique, et une discussion encore plus obscure avec PRINGSHEIM ⁽²⁾ (qui prétendait le continu fonctionnel analogue à un continu numérique pluridimensionnel). La seule justification de cette affirmation se trouverait dans les travaux contemporains ou postérieurs de

(1) DU BOIS REYMOND (VII).

(2) PRINGSHEIM (I).

STOLZ ⁽¹⁾, BETTAZZI ⁽²⁾, HAHN ⁽³⁾ sur les systèmes non archimédiens : ceux-ci peuvent, en effet, toujours se représenter par des nombres complexes dont les unités, en général en nombre infini, forment une suite ordonnée. Mais le but du calcul infinitaire est de comparer les uns aux autres les degrés de croissance : la pluridimensionalité serait absurde.

BOREL clarifie le débat en montrant ⁽⁴⁾, d'une part, que l'ensemble des ordres de croissance, ayant même puissance que celui des fonctions continues, a la puissance du continu linéaire ; d'autre part, que la comparaison avec celui-ci ne vaut que si on part d'une échelle donnée de types croissants et si on se borne à la compléter uniquement par multiplication de fonctions qui y figurent, exactement comme on obtient les nombres rationnels à partir des entiers. Dans ce cas il sera même possible de définir des coupures, et une *fonction idéale* correspondante. DU BOIS REYMOND en avait déjà eu l'idée à propos de l'intégrale de FOURIER : il s'agissait d'examiner la convergence d'intégrales de la forme

$$I = \int_{\varepsilon}^{\alpha} \frac{d\alpha}{\varphi(\alpha)} \text{ lorsque } \varepsilon \text{ tend vers } 0.$$

On voit que si on a :

$$\varphi(\alpha) \sim \frac{I}{l \frac{I}{\alpha} \cdot l_2 \frac{I}{\alpha} \dots l_r \frac{I}{\alpha}} \quad \lim I = \infty$$

Si, au contraire,

$$\varphi(\alpha) \sim \frac{I}{l \frac{I}{\alpha} \cdot l_2 \frac{I}{\alpha} \dots \left(l_r \frac{I}{\alpha} \right)^{1+\mu}} \quad \lim I \text{ est finie}$$

(1) STOLZ (I).

(2) BETTAZZI (I).

(3) HAHN (I).

(4) BOREL (I), p. 117-119.

(μ étant un nombre positif quelconque). La coupure entre les degrés infinitaires faisant diverger et converger l'intégrale, définira la fonction idéale $\tau(\alpha)$. On démontre qu'il est impossible que $\tau(\alpha)$ soit de même ordre infintaire qu'une fonction logarithmique quelconque. Pour $\varphi(\alpha) = \tau(\alpha)$, l'intégrale ne diverge ni ne converge, «représentation choquante pour le sens commun ». Mais « le miracle n'est pas plus grand... que d'admettre de nouveaux nombres entre les nombres rationnels pourtant denses à volonté » (1).

Ainsi échoue à demi cette première tentative d'application de la théorie des ensembles à un problème mathématique. L'ambition de DU BOIS REYMOND était d'embrasser tout le système des ordres de croissance dans un seul « continu » d'ordre nouveau. Mais il aurait fallu posséder des moyens de caractériser celui-ci. Or la notion claire du continu linéaire n'était même pas encore dégagée : le procédé de la diagonale seul se trouvait insuffisant. D'autres instruments propres à la théorie des ensembles étaient nécessaires, que justement à cette époque, dans un tout autre domaine, CANTOR et DEDEKIND travaillent à mettre au point.

(1) DU BOIS REYMOND (IX), p. 165.

CHAPITRE II

La création cantorienne.

Deux écrits méritent de figurer avant ceux de CANTOR dans une histoire de la théorie parce qu'essayant les premiers de soumettre à un traitement systématique les ensembles en tant que tels : les *Paradoxes de l'Infini* de BOLZANO et un fragment de DEDEKIND, non publié de son vivant : *Propositions générales sur des espaces*.

« BOLZANO, écrit CANTOR en 1883, est peut-être le seul chez lequel les nombres proprement infinis jouissent d'un certain droit de cité ; au moins en parle-t-il à plusieurs reprises ⁽¹⁾. » C'est par l'isolement de la notion d'ensemble que BOLZANO peut, en effet, distinguer de l'infini improprement dit — l'infini de l'analyse qui ne concerne qu'une possibilité de variation et représente des nombres finis pris aussi grands qu'on veut — le véritable infini, sur lequel jusque là rien que de vague ou d'absurde n'avait été dit. Outre la définition générale d'un ensemble, le livre apporte la notion de correspondance biunivoque ; une ébauche d'arithmétique des ensembles, un essai de définition abstraite du continu. « Un certain système (INBEGRIFF) de choses est un tout consistant en certaines parties... Un système dans lequel l'ordre des parties est indifférent s'appelle un ensemble ⁽²⁾. » La notion de parties est encore vague : si ce sont « les unités d'une certaine espèce $A...$, c'est-à-dire des objets subsumés sous un concept A » l'ensemble est dit une

(1) CANTOR (XV), p. 180.

(2) BOLZANO (II), p. 4.

« multiplicité » ⁽¹⁾ (VIELHEIT). Ainsi l'ensemble est considéré comme une réalité existante en soi, indépendamment de nos moyens de l'atteindre. Il suffit qu'on puisse décider si un élément donné lui appartient ou non. « Nous avons déterminé de la façon la plus complète l'ensemble... infini des points entre m et n , aussitôt que nous avons déterminé les deux points m et n . Car alors, par ce peu de mots, il est déjà décidé exactement pour tout autre point s'il appartient ou non à cet ensemble ⁽²⁾. » D'où la possibilité qu'un ensemble soit infini en acte, c'est-à-dire que nous ne puissions l'embrasser du regard et même que nous ne sachions l'engendrer. Des relations entre de tels ensembles — un calcul de l'infini — peuvent être établies grâce d'abord à la mise en correspondance biunivoque : « deux ensembles sont dans un tel rapport l'un avec l'autre que pour toute partie a du premier A , nous pouvons, d'après une certaine règle, trouver une partie b du second B et que toutes les paires $(a + b)$ ainsi formées contiennent toute chose qui se trouve dans A ou dans B et chacune seulement une fois » ⁽³⁾. L'addition et la multiplication (contrairement à ce que croyait CANTOR) ⁽⁴⁾ sont correctement définies : « il peut y avoir des systèmes dont les parties... sont elles-mêmes des systèmes. Parmi eux tels systèmes que rien en eux d'essentiel n'est changé si on considère les parties des parties comme parties du tout. Je les appelle... sommes » ⁽⁵⁾. On n'introduit la multiplication que pour l'ensemble des entiers, représenté par la notation :

(1) *Ibid.*

(2) *Ibid.*, p. 38.

(3) P. 31.

(4) « Je ne m'accorde en rien avec la façon dont il les traite [les nombres infinis], sans pouvoir donner d'eux une définition correcte et je considère par exemple les §§ 29-33 [où se trouve la définition du produit] comme inconsistants et erronés ». CANTOR (XV), p. 180.

(5) *Ibid.*, p. 4.

$$\overset{\circ}{N} = 1^{\circ} + 2^{\circ} + 3^{\circ} + \dots$$

(les $^{\circ}$ et les $+$ indiquent qu'il ne s'agit pas d'une somme arithmétique), le produit de cet ensemble par lui-même sera :

$$(\overset{\circ}{N})^2 = 1^{\circ} \cdot \overset{\circ}{N} + 2^{\circ} \cdot \overset{\circ}{N} + 3^{\circ} \cdot \overset{\circ}{N} + \dots$$

à chaque élément du multiplicateur est substitué un ensemble équivalent au multiplicande ⁽¹⁾. Ici, comme de façon générale, il n'y a pas dégagement suffisant du procédé par rapport aux exemples où il est saisi intuitivement. D'où un flottement à propos d'abord de la notion de partie — sous ensemble — et de l'égalité de deux ensembles. Les prédicats *fini* et *infini* ne sont déterminés que pour les multiplicités (VIELHEITEN) dont on connaît la loi d'engendrement, qui constituent des séries : lorsque la multiplicité peut être obtenue tout entière par adjonction successive d'un élément, elle est finie ; « une multiplicité telle que tout ensemble fini en représente une partie est appelée multiplicité infinie. » ⁽²⁾ Quoique le premier exemple d'ensemble infini ne soit pas emprunté aux mathématiques — ensemble de toutes les propositions — l'image de la suite des nombres (avec un ordre) est dominante ; la mise en correspondance biunivoque est ici masquée par le processus de numération. Bien plus, le paradoxe des ensembles infinis — que DEDEKIND transformera en définition de ces ensembles — amène à diminuer son importance, là où elle est expressément utilisée : on peut mettre en correspondance biunivoque les éléments d'un ensemble et ceux d'une de ses parties. Donc l'égalité de deux ensembles devra être définie autrement : « une autre considération doit intervenir..., savoir que les deux ensembles ont mêmes principes de détermination, par

(1) P. 45.

(2) P. 6.

exemple le même mode d'engendrement (1) ». Ici encore c'est la référence aux exemples qui fait faire fausse route : les deux ensembles A (intervalle de nombres 0...5) et B (intervalle 0...12) ne sont pas égaux parce que « la distance que deux grandeurs ont entre elles dans B est autre (plus grande) que la distance qui sépare dans A les deux grandeurs correspondantes ». Mais comment avoir une théorie autonome si une définition des structures empruntée à d'autres théories doit intervenir dès la relation d'égalité ?

A propos du continu, même flottement. Le but est abstrait : il s'agit de caractériser le continu comme ensemble infini ayant un « certain ordre » établi entre ses éléments : « il n'y a de continu que là où se trouve un système d'objets simples (de points dans le temps ou dans l'espace, ou encore de substances) placés de telle sorte que chacun d'eux possède, à toute distance si petite soit-elle, un voisin appartenant au système » (2). La deuxième partie de la définition la restreint aux espaces métriques — et ne laisse aucune place à des considérations de voisinage : il ne s'agit pas de densité en soi, comme on a cru depuis (3). On le voit par l'exemple du point 0, *isolé* dans le segment 0...1 qui n'est donc plus un continu si on garde 0 et si on en retranche les points d'abscisse $\frac{1}{2^n}$, parce que « pour toutes les distances de forme $\frac{1}{2^n}$ il n'y a pas de voisin pour 0 ». La définition n'est donc qu'un premier énoncé de l'axiome de CANTOR (4). — La dimension est définie par récurrence : « un étendu (c'est-à-dire un continu) est *simplement étendu* ou ligne, si tout point, pour toute distance si petite soit-elle, y a un ou plusieurs

(1) P. 31.

(2) P. 73.

(3) Le voisin ne doit pas se trouver à l'intérieur d'un système de rayon quelconque mais *sur la surface* de cette sphère.

(4) A tout nombre réel correspond sur une droite un point ayant ce nombre pour abscisse.

voisins, mais jamais assez pour que leur système à lui seul constitue déjà un étendu », il a deux dimensions « si tout point pour toute distance... a une ligne entière de points comme voisins, etc... (1) ». Enfin on dit qu'un étendu est « *d'un seul tenant* s'il n'a aucune partie formant elle-même un étendu dans laquelle il n'y ait au moins un point qui, compté aux autres parties, ne forme avec elles également un étendu (2) ».

L'intérêt des dernières définitions est réduit par l'étroitesse de celle du continu. Le fragment de DEDEKIND (3) marque un progrès en introduisant la notion de voisinage (défini métriquement) : il s'agissait, à l'occasion des leçons de DEDEKIND sur le potentiel, et, « pour fonder plus rigoureusement le principe de DIRICHLET », de développer une *théorie des domaines* (GEBIETSLEHRE) « *ab ovo* sans faire appel à l'intuition géométrique (4) ». On définit ensembles ouverts, points intérieurs, extérieurs et frontières : « un système P de points $p, p' \dots$ forme un *corps* [ensemble ouvert] si pour tout point p on peut donner une longueur δ telle que tous les points distants de p de moins de δ appartiennent à P . Les points $p, p' \dots$ sont à l'intérieur de P » (5). De même « si P est un corps et m un point autour duquel on peut décrire une sphère dont tous les points sont non-intérieurs à P , le point m est à l'extérieur de P ». Enfin la *frontière* d'un corps est l'ensemble des points ni extérieurs ni intérieurs. — Quelques théorèmes : s'il y a un point extérieur à un corps, il y en a une infinité et leur système forme un corps ; le système des points frontières d'un corps ne peut consti-

(1) P. 80, n.

(2) P. 95.

(3) DEDEKIND (I), t. II, p. 353-355.

(4) L. à CANTOR, 19 janvier 1879 : « je ne me permettrais pas un tel jugement, si il y a de nombreuses années lorsque je voulais encore éditer la théorie de DIRICHLET sur le potentiel... je ne m'étais déjà assez occupé de ces questions ». CANTOR-DEDEKIND (I), p. 47.

(5) DEDEKIND (I), t. II, p. 353.

tuer un corps. « Pris par la mise au point de la théorie des nombres de DIRICHLET », DEDEKIND abandonne ces définitions qui, en 1879, lui semblent encore « constituer une bonne base » : quoique sans prolongement — au moins immédiat — elles sont une sorte de préface aux travaux de DEDEKIND sur les ensembles, échelonnés parallèlement à ceux de CANTOR à partir de 1872. L'échange entre les deux mathématiciens qui se poursuit tout au long de la création de CANTOR les situe donc logiquement avant celle-ci.

* * *

De 1873 à 1884 — jusqu'aux mémoires *Sur les ensembles infinis et linéaires de points*, premier exposé en forme de la théorie — la création cantorienne se caractérise par son intime subordination à l'analyse. C'est d'ensembles de points ou de nombres — non d'ensembles abstraits — qu'il s'agit, ce sont les applications analytiques qui importent, et si de nouveaux êtres sont introduits, avec une arithmétique qui leur est propre, leur justification est que seuls ils permettent « d'aller plus loin ⁽¹⁾ » dans cette branche de la théorie des fonctions qu'est l'étude des ensembles de points. D'où l'enrobement des méthodes nouvelles dans une gaine intuitive sous laquelle il faut chercher leur véritable portée ; mais aussi dans la fraîcheur de leur apparition une sorte de naïveté qui laisse apparents les liens, permet de suivre la continuité d'un mouvement nécessaire. La création se fait en deux étapes : de 1873 à 1877, découverte et distinction de deux puissances ; de 1878 à 1884, exploration du continu, établissement de la théorie avec, comme complément attardé, en 1890, l'article où apparaît pour la première fois le procédé de la diagonale.

La première période n'est guère que préparation, mais

(1) CANTOR (XV), p. 165.

décisive. BOLZANO avait abouti à une impasse : pour explorer l'infini, l'application ⁽¹⁾, seul procédé précis (généralisation de la numération), se révélait impropre puisque confondant tout et partie. Avec les deux résultats de CANTOR l'application est réhabilitée, un nouvel infini conquis au raisonnement mathématique : non seulement négativement on peut démontrer qu'il y a plus grand que l'ensemble des entiers, mais positivement, quoique l'on dépasse toute loi de série, on peut égaler deux espèces distinctes de ce plus grand : la mathématique sort cette fois de l'indéfini, une théorie de l'infini est fondée. Le point initial des deux recherches se trouve dans la définition des nombres irrationnels : d'une part, représentation du continu comme ensemble de toutes les limites superposées de suites de nombres rationnels ; d'autre part, emploi systématique de la réduction de suites à *n^{uple}* entrée en suites simples ou *linéarisation*, — procédé qui dans les deux cas ouvre le problème.

En 1873, CANTOR se demande si l'ensemble des nombres réels ne peut être mis sous forme de suite simple : « à première vue on se dit non, c'est impossible, car (n) [l'ensemble des entiers] est formé de parties discrètes, tandis que (x) [l'ensemble des nombres réels] constitue un continu ; seulement cette objection n'avance à rien... Ne serait-on pas aussi enclin, à première vue, à affirmer que (n) ne se laisse pas coordonner de façon biunivoque au système $\left(\frac{p}{q}\right)$ de tous les nombres rationnels positifs $\frac{p}{q}$? ⁽²⁾. » Cependant, non seulement l'ensemble des nombres rationnels, mais, comme le montre DEDEKIND ⁽³⁾, l'ensemble des nombres algébriques se laisse mettre sous forme de suite simple et, de façon plus générale, tout ensemble ayant pour éléments $a_{n_1, n_2, \dots, n_i}$, dont les indices en nombre fini parcourent indépendamment les

(1) Nous dirons désormais *application* (ABBILDUNG) d'un ensemble sur un autre au lieu de mise en correspondance *biunivoque* de leurs éléments.

(2) L. à DEDEKIND, 29 novembre 1873. CANTOR-DEDEKIND (I), p. 13.

(3) CANTOR-DEDEKIND (I), p. 18.

uns des autres la suite des entiers. L'intérêt de la question, ajoute CANTOR comme pour se justifier, est qu'une réponse négative « donnerait une nouvelle démonstration au théorème de LIOUVILLE qu'il existe des nombres transcendants » (1). Il s'agit donc bien encore d'analyse. Le 7 décembre la réponse est trouvée grâce au procédé de l'emboîtement des intervalles, déduit directement de deux des propriétés du continu : entre deux éléments il y en a toujours d'autres (densité) ; toute suite convergente d'éléments définit un élément du système (compacité en soi). De la 1^{re} propriété découle la présence entre deux termes de la suite :

$$(I) \qquad a_1, a_2 \dots a_n \dots$$

— où seraient rangés les nombres réels — d'autres termes de (I), et ainsi indéfiniment : constitution d'une suite d'intervalles emboîtés tels que tous les termes de (I) dont le rang est inférieur à celui des extrémités d'un intervalle soient extérieurs à cet intervalle. De la 2^e propriété suit qu'il y a au moins un point m intérieur à tous les intervalles, le point limite commun des deux suites d'extrémités, si la longueur des intervalles tend vers 0 : m ne peut donc appartenir à (I).

La linéarisation n'intervient que négativement ; le théorème montre qu'elle est inefficace devant le continu : on ne peut pas coordonner (2) à (n) le système $(a_{n_1, n_2} \dots)$ où les indices sont en nombre infini. Immédiatement — avec la même attente de réponse négative — se repose le problème pour le passage du continu simple au continu à n dimensions. « Dans le même ordre d'idées, écrit CANTOR le 5 janvier 1874, se présente aussi la question suivante : peut-on coordonner de façon... [biunivoque]... une surface et une ligne ?... Il me semble encore pour l'instant — quoique

(1) L. à DEDEKIND, 2 décembre 1873, *ibid.*, p. 13.

(2) Nous représentons, comme fait CANTOR, une suite dont le terme général est n par (n) .

l'on se sente tellement contraint ici au *non*, que l'on voudrait tenir une démonstration presque pour superflue — que la réponse présente de grandes difficultés ⁽¹⁾. » Absurde question, déclarent les mathématiciens de Berlin, car « il va de soi que deux variables indépendantes ne peuvent se laisser ramener à une ⁽²⁾ ». Pourtant trois ans après, la linéarisation l'emporte ⁽³⁾ : un système de n fractions décimales indéfinies (les n coordonnées du point dans l'espace à n dimensions) peut se mettre sous forme d'une seule fraction décimale indéfinie (abscisse du point de la droite). Mais il faut que le procédé donne une correspondance biunivoque et exhaustive, difficulté signalée par DEDEKIND : si on exclut les fractions terminées par une suite infinie de 0 (pour obtenir l'unicité de la représentation décimale) les points de la droite dont l'abscisse comporte un développement avec périodicité de 0 (de période inférieure ou égale à n) n'auront pas de correspondant dans l'espace à n dimensions ⁽⁴⁾.

(I) CANTOR-DEDEKIND (I), p. 20.

(2) L. à DEDEKIND, 18 mai 1874, *ibid.*, p. 21.

(3) L. à DEDEKIND, 20 juin 1877, *ibid.*, p. 25 : si on a le système :

[illegible]

on en tire : $y = \beta_1 \frac{1}{IO} + \beta_2 \frac{1}{IO^2} + \dots + \beta_v \frac{1}{IO^v} + \dots$ en posant :

$$\beta(n-1)\rho + \sigma = \alpha_{\sigma,n}.$$

(4) L. DEDEKIND à CANTOR, 22 juin 1877, *ibid.*, p. 27. Si on prend pour $n = 2$:

$$x_1 = \frac{\alpha_1}{10} + \frac{\alpha_2}{10^2} + \dots$$
$$x_2 = \frac{\beta_1}{10} + \frac{\beta_2}{10^2} + \dots$$

$$y = \frac{\gamma_1}{10} + \frac{\gamma_2}{10^2} + \dots \text{ avec } \gamma_1 = \alpha_1 \quad \gamma_2 = \beta_1 \quad \gamma_3 = \alpha_2 \quad \gamma_4 = \beta_2, \text{ etc.}$$

des fractions comme $y = 0,310507090_6 0\alpha_7 0\alpha_8 0\dots$ ne seront pas représentées puisqu'on aurait $x_2 = \frac{3}{10} + 0 + 0 \dots$ ce qui est exclu par convention.

Dans une nouvelle démonstration, le procédé de linéarisation est appliqué seulement aux systèmes de nombres irrationnels définis sans ambiguïté par leur développement en fraction continue. Mais qu'ajoute aux infinis en présence le système des nombres rationnels ? ici doit porter l'effort, uniquement d'ailleurs pour dégager les notions, car aucune méthode nouvelle n'est nécessaire. CANTOR avance avec timidité, en référence constante à l'intuition, comme pour se rassurer devant le résultat : « je le vois mais je ne le crois pas » (1). Trois étapes :

1° l'ensemble A , continu linéaire privé des nombres rationnels est appliqué sur l'ensemble B , continu linéaire privé d'une suite croissante (ϵ_v) de nombres irrationnels : l'équivalence est démontrée entre deux variables qui parcourent A et B (et non pas directement entre A et B) et dont on voit qu'on peut faire correspondre les valeurs ;

2° l'ensemble B est appliqué sur le continu linéaire complet M par utilisation itérée, pour chaque intervalle $(\epsilon_v \dots \epsilon_{v+1})$, du théorème que « un nombre y qui peut prendre toutes les valeurs de l'intervalle $(0 \dots 1)$ à l'unique

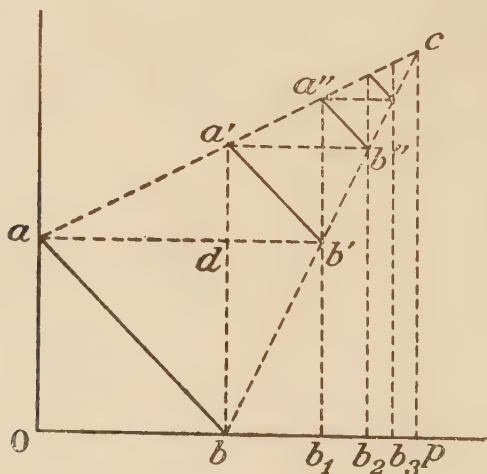


FIG. 2.

(1) L. à DEDEKIND, 29 juin 1877, *ibid.*, p. 34.

exception de la valeur 0, peut être coordonné de façon univoque à un nombre x qui prend toutes les valeurs de $(0...1)$ sans exception ». Ici la démonstration n'est même plus analytique mais graphique, grâce à la « courbe remarquable » ⁽¹⁾ à laquelle CANTOR attache tant d'importance qu'il la reproduit de sa lettre à DEDEKIND à son article du *Journal de Crelle*, alors que figure dans celui-ci une autre démonstration complète qui tient en quelques lignes ;

3^o enfin, puisque tout continu linéaire complet M est applicable sur l'ensemble linéaire irrationnel A , le continu à n dimensions M_n , défini par n continus linéaires, s'appliquera sur l'ensemble irrationnel à n dimensions A_n donc, par l'intermédiaire de l'ensemble A , sur le continu linéaire M .

Enfantement pénible de la notion de puissance qu'à la fois la transitivité des applications et l'étude des rapports entre ensemble et sous-ensemble obligent à dégager complètement. « Si deux ensembles bien définis M et N se laissent coordonner l'un à l'autre, élément par élément, de façon univoque et complète, je me sers de l'expression... qu'ils ont *égale puissance* ou qu'ils sont équivalents. Par une *partie intégrante* (BESTANDTHEIL) d'un ensemble M nous entendons tout autre ensemble M' dont les éléments sont en même temps éléments de M . Si deux ensembles M et N ne sont pas d'égale puissance, ou bien M a une puissance égale à celle d'une partie intégrante de N , ou N celle d'une partie intégrante de M ; dans le premier cas nous nommons la puissance de M *plus petite*, dans le deuxième cas nous la nommons *plus grande* que la puissance de N ⁽²⁾. » Formule (et

(1) La courbe y est constituée par les segments $ab, a'b', a''b''...$ et le point c , en sont exclus les points $b, b', b^{(v)}...$: on voit que x prend toutes les valeurs de 0 à 1, au contraire y toutes les valeurs de 0 à 1 sauf 0.

(2) CANTOR (VIII), p. 119. Cf. plus tard : « j'ai emprunté le mot *puissance* à J. STEINER qui l'emploie dans un sens tout spécial, cependant voisin pour exprimer que deux figures sont coordonnées l'une à l'autre *projectivement* de telle sorte qu'à tout élément de l'une correspond un et un seul élément de l'autre ; dans le concept de puissance ici employé on conserve la corres-

solution supposée) déjà du problème de la comparabilité. Ainsi un calcul des puissances se trouve ébauché. La puissance est un nombre : elle se confond avec le nombre entier positif pour les ensembles finis « puisque en effet deux tels ensembles ont et n'ont puissance égale que si le nombre de leurs éléments est le même ». Pour les deux puissances de l'ensemble des entiers et du continu, non seulement la relation d'inégalité est démontrée, mais le résultat d'additions et de multiplications :

1^o « Si M' , M'' , M''' est une suite finie ou infinie d'ensembles ayant chacun la puissance de la suite des entiers positifs, l'ensemble M formé de la réunion de ces ensembles a aussi la même puissance (1). »

2^o L'addition de la puissance des entiers à la puissance du continu (ou la soustraction de la 1^{re} de la 2^e) donne la puissance du continu (parties 1 et 2 du théorème).

3^o La multiplication de la puissance du continu par elle-même donne la puissance du continu. C'était au moins un des aspects sous lesquels se présentait le problème : « comme si l'infini des points d'une surface était obtenu à partir des points d'une ligne par élévation au carré, celui des points d'un volume par élévation au cube (2) ». Toutefois ce calcul — en particulier le dernier résultat — n'est pas encore expressément dégagé : pour que les puissances deviennent effectivement des nombres, il faut que leur mode d'engendrement, en particulier d'abord le passage du dénombrable au continu, soit découvert. C'est donc l'analyse du continu qui passe au premier plan, avec d'ailleurs son autorité concrète : après avoir en partie détruit la notion intuitive (pour la dimension) il s'agit de reconstruire. Mais dans la joie d'avoir dominé « cette force étonnante qui réside dans

pondance biunivoque, mais en revanche n'est imposée aucune restriction à la loi de coordination ». CANTOR (XIII), p. 151.

(1) CANTOR (VIII), p. 120.

(2) L. à DEDEKIND, 20 juin 1877. CANTOR-DEDEKIND (I), p. 25.

les nombres réels rationnels et irrationnels, capable de déterminer les éléments d'un système continu à ρ dimensions [ou à une infinité de dimensions, comme le montre une extension immédiatement faite de la linéarisation] de façon univoque à l'aide d'une seule coordonnée » ⁽¹⁾, disparaît toute timidité devant l'infini : « tous les ensembles linéaires imaginables de points ou de nombres réels » sont répartis en classes déterminées d'après leur puissance et déjà se formule le problème central : « par un procédé d'induction — qui ne sera pas précisé ici — est approché le théorème que le nombre des classes donné par cette répartition est fini et égal à deux ⁽²⁾ ».

Ainsi s'oriente définitivement la théorie : parallèlement aux recherches de topologie proprement dite — sur la dimension et la connexion — s'ouvre une série de travaux (théorie des ensembles de points mêlée avec théorie abstraite des ensembles) qui n'ont dès maintenant d'autre but que de prouver la validité de l'hypothèse et d'établir une relation entre les deux puissances. Le moyen d'action est emprunté au dernier mémoire *Sur les séries trigonométriques* dont l'influence souterraine semble n'avoir pas cessé ⁽³⁾ : c'est l'opération de la dérivation, opération centrale, — que U. DINI a utilisée « dans son dernier ouvrage pour une série de remarquables généralisations de théorèmes connus d'analyse... Sur cette opération, comme nous le montrerons plus tard, est fondée la plus simple et en même temps la plus complète explication et détermination du continu ⁽⁴⁾ ». C'est d'elle en effet que proviennent également les deux espèces d'objets essentiels à l'édifice cantorien et grâce aux-

(1) L. 25 juin 1877, *ibid.*, p. 34.

(2) CANTOR (VIII), p. 132.

(3) « Pour ce qui concerne les nombres ordinaux transfinis il me paraît vraisemblable que j'en avais déjà une représentation en 1871 ». L. à JOURDAIN, 31 août 1905 in JOURDAIN (I), p. 28.

(4) CANTOR (X), p. 139.

quels le problème sera attaqué sinon résolu, les nombres transfinis et les ensembles fermés et parfaits.

En 1871, au moins dans le mémoire publié, on ne distinguait que deux *genres* (GATTUNGEN) d'ensembles : ceux qu'un nombre fini de dérivations réduisaient à 0 ; ceux « pour lesquels la suite indéfinie des dérivés ne s'arrête pas ». En 1877, dans la *classe* des ensembles de 1^{re} puissance, figure le 1^{er} *genre* d'ensembles. Mais comment se comportent, du point de vue de la puissance, les ensembles de 2^e genre ? A une première confrontation de ces deux modes de répartition sont consacrés les mémoires échelonnés entre 1879 et 1882.

D'abord deux nouvelles notions qui serviront d'instruments : la densité et le recouvrement par des intervalles. La densité avait été définie en 1872 par DEDEKIND en tant que propriété commune aux systèmes des nombres rationnels et réels. CANTOR marque seulement sa parenté avec l'opération de la dérivation : être dense dans un intervalle A c'est, pour un ensemble, posséder des points dans tout intervalle déterminé au sein de A : donc le premier dérivé d'un ensemble dense coïncide avec l'intervalle tout entier. Un ensemble dense est nécessairement de 2^e genre (la réciproque n'étant pas vraie). Un ensemble partout non dense est tel que dans tout intervalle on puisse trouver un intervalle où il ne possède aucun point. La notion de recouvrement par intervalles est empruntée à DU BOIS REYMOND : une fonction discontinue peut être intégrée si l'ensemble de ses points de discontinuité forme un *groupe intégrable*, c'est-à-dire peut être recouvert par une suite finie d'intervalles dont la longueur totale est nulle. Un tel ensemble est évidemment partout non dense. Le résultat important obtenu par CANTOR est que tout ensemble P dont le 1^{er} dérivé P' est dénombrable forme un groupe intégrable : en effet, ni cet ensemble, ni son dérivé, ni la somme $P + P'$ ne peuvent être denses dans aucun intervalle ; chaque point ex-

térieur à $P + P'$ peut donc être entouré d'un intervalle dont aucun point sauf les extrémités n'appartient à $P + P'$; on démontre que la somme de ces intervalles *contigus* à $P + P'$ est égale à la longueur totale de l'intervalle global (dans lequel P et P' sont définis), en utilisant une propriété élémentaire des fonctions continues ⁽¹⁾ ; il sera donc possible d'enfermer $P + P'$ dans un nombre fini d'intervalles de longueur arbitrairement petite. Le raisonnement par considération des intervalles contigus est nouveau (dû en partie à DU BOIS REYMOND) et resservira pour les ensembles parfaits.

En second lieu précision et extension des notions déjà acquises : la dérivation est généralisée au cas de l'espace à n dimensions ; on démontre, en transformant par rayons vecteurs réciproques, l'espace à n dimensions en une surface (d'aire finie) à $n + 1$ dimensions, et en utilisant ensuite le procédé de l'*échelle dénombrable* (il n'y a qu'un nombre fini d'intervalles d'aire supérieure à un nombre rationnel donné puisque la somme totale est finie), qu'il n'y a qu'une infinité dénombrable d'intervalles (à n dimensions) non empiétants, — théorème nécessaire pour la suite. La puissance est définie pour des ensembles non plus de points, mais d'objets quelconques « lignes, surfaces, corps ». Pour tout couple d'ensembles bien définis « on doit considérer comme intrinsèquement déterminé s'ils ont ou non même puissance ». Un ensemble d'éléments est « *bien défini*, si sur la

(1) La fin du raisonnement peut être ainsi donnée : soit $f(x) = x - \sigma$ σ étant la somme des intervalles contigus entre 0 et x ($P + P'$ est supposé situé sur l'intervalle 0...1) :

$$\begin{array}{c} 0 \qquad \qquad \qquad 1 \\ | \text{-----} | \end{array}$$

$f(x)$ est continue car on a évidemment :

$$|f(x+L) - f(x)| \leq L.$$

Soit l la longueur totale, des intervalles contigus : $f(x)$ varie de 0 à $1-l$; or, à l'intérieur d'un intervalle contigu $f(x)$, est constante, comme $P + P'$ est dénombrable elle ne prend donc qu'une infinité de valeurs, comme elle est continue elle prend toutes les valeurs entre 0 et $1-l$, donc $l = 1$.

base de sa définition, et en conséquence du principe logique du tiers exclu, on doit considérer comme intrinsèquement déterminé, aussi bien si un objet quelconque appartenant à la même sphère conceptuelle fait partie ou non de l'ensemble, qu'également si deux objets appartenant à l'ensemble sont identiques malgré les différentes façons dont ils sont donnés ⁽¹⁾ ». De ce processus d'abstraction et de projection en quelque sorte dans un domaine objectif, le problème du continu reçoit une acuité plus grande : il doit être intrinsèquement déterminé si les ensembles de points ne peuvent avoir que les deux puissances connues ; d'autre part une définition abstraite de ces puissances et de leur rapport est possible si les deux types d'ensembles auxquels elles s'appliquent peuvent être suffisamment caractérisés conceptuellement. La dérivation va-t-elle permettre cette caractérisation ? L'analyse mécanique, nécessaire en quelque sorte, qu'elle produit dans un ensemble le fait espérer : on sait déjà qu'un ensemble dont le $n^{\text{ième}}$ dérivé est nul a la première puissance ; mais on peut aller plus loin : il suffit de remarquer qu'à partir du 2^{e} dérivé tous les dérivés sont partie intégrante de leur prédécesseur. Si donc tous les dérivés finis ont des points, on considérera comme dérivé de rang ∞, p^∞ , l'ensemble des points communs à toute la suite. En dérivant à nouveau on introduit les symboles $\infty + 1, \infty + 2, \dots, \infty + n$, puis une nouvelle partie commune donne $p^2 \infty, \dots p^{\infty^2}, \dots p^{\infty \infty}, p^{\infty \infty \infty} \dots$ « Nous voyons ici une génération dialectique de concepts qui conduit toujours plus loin, et libre de tout arbitraire reste nécessaire en soi et conséquente ⁽²⁾. »

Grâce à elle, en effet, sont procurés immédiatement au sujet de la puissance trois résultats d'extension croissante :

1^o Si l'on appelle isolé un ensemble qui n'a pas de point commun avec son dérivé (par exemple la différence $P' - P''$

(1) CANTOR (XIII), p. 150.

(2) CANTOR (XI), p. 148.

entre le 1^{er} et le 2^e dérivé), on remarque que les dérivations successives à partir de la 2^e ont pour effet de retrancher un ensemble isolé. Or tout ensemble isolé est dénombrable : chacun de ses points peut être, en effet, enfermé dans un intervalle (à n dimensions si l'ensemble est situé dans un espace à n dimensions) non empiétant sur les autres et comme on a vu qu'il n'y a qu'une infinité dénombrable de tels intervalles l'ensemble est dénombrable.

2^o Si le 1^{er} dérivé P' d'un ensemble P est dénombrable, P est dénombrable : en effet, un ensemble se compose d'un ensemble isolé et de la partie commune avec son dérivé.

3^o Tout ensemble dont le dérivé d'ordre α (α étant soit un nombre fini, soit un quelconque des précédents symboles transfinis) est nul et dénombrable. La démonstration est immédiate lorsque α est fini. Lorsque α est un symbole transfini, CANTOR déclare « qu'elle doit se faire exactement de même, à l'aide de l'induction complète ⁽¹⁾ », sans même remarquer l'extension donnée du même coup au procédé d'induction. C'est que le mouvement dialectique engendrant la suite des ∞ lui apparaît avec l'évidence contraignante de la logique. Mais ici se manifeste aussi sa limitation : l'induction complète ne peut prouver la dénombrabilité qu'à condition de ne comporter elle-même qu'une infinité dénombrable d'étapes. Si l'adhérence reste entre induction et génération des ∞ , impossible de dépasser le dénombrable, la liaison avec le continu est manquée. Si l'on veut atteindre celle-ci il faut que les symboles ∞ soient autre chose que les étapes d'une génération progressive, mais les représentants d'une réalité immanente que des moyens nouveaux permettent d'atteindre, il faut s'engager dans le domaine véritable du transfini.

Tel est le dilemme devant lequel se trouvait CANTOR à la

(1) CANTOR (XIV), p. 160.

fin de l'été 1882 : jusque-là rien que de normal dans l'édifice construit ; si les résultats sont étonnants, les méthodes employées sont celles de l'analyse ordinaire, aucun mathématicien, si empiriste soit-il, n'y peut rien trouver à redire. Dans le mémoire de mars 1882 ⁽¹⁾ le ton a sans doute déjà changé : mais l'affirmation d'un en soi des vérités mathématiques n'a pas de conséquence pratique. La décision effective pour l'égalité de puissance de deux ensembles est comparée à la décision pour savoir si un nombre donné est ou non transcendant, « un des plus difficiles problèmes » de l'analyse ⁽²⁾. C'est plutôt l'idée que la théorie des ensembles domine et logiquement précède le reste des mathématiques qui pourrait annoncer un changement de point de vue : « la théorie des ensembles, dans la conception ici exposée, embrasse — si nous ne considérons que le mathématique et laissons provisoirement de côté les autres sphères conceptuelles — les domaines de l'arithmétique, de la théorie des fonctions et de la géométrie ; elle les réunit sur la base du concept de puissance en une unité supérieure. *Discontinu* et *continu* se trouvent de la sorte considérés du même point de vue et mesurés avec une mesure commune ⁽³⁾ ». Quels peuvent être cette mesure, ce point de vue communs, sinon ceux que procure une refonte de la notion de nombre ? La phrase étrange sur « les autres sphères conceptuelles » semble se référer aux enquêtes philosophiques sur l'infini, qui tiendront la première place dans le mémoire d'octobre. Toutefois mathématiquement rien encore n'apparaît : dans la correspondance avec DEDEKIND jusqu'aux premiers jours d'octobre il n'est question que de la définition topologique du continu. C'est, semble-t-il, dans l'exaspération de l'échec, après l'inefficacité reconnue de la dérivation, que CANTOR se

(1) CANTOR (XIII) daté 31 mars 1882.

(2) *Ibid.*, p. 150.

(3) P. 152.

décide à franchir le pas : dans une soudaine illumination, il construit, en quelques jours sans doute, le mémoire d'octobre, il se lance à corps perdu dans l'infini actuel : « Justement depuis nos dernières rencontres à Harzburg et Eisenach, écrit-il le 5 novembre à DEDEKIND ⁽¹⁾, Dieu tout-puissant a voulu que je parvienne aux lumières les plus étonnantes et les plus inattendues sur la théorie des ensembles et sur la théorie des nombres, ou, bien plus, que je trouve ce qui depuis des années a fermenté en moi, ce vers quoi j'ai fait de longues recherches ⁽²⁾. » L'exaltation du mémoire lui-même, ce désordre de considérations mathématiques pêle-mêle avec les références philosophiques les plus confuses — la règle d'addition des ordinaux $\omega + 1 > \omega$ est présentée comme solution au problème spinoziste du rapport entre modes infinis et modes finis, — aussi bien que les excuses réitérées dans l'article et dans la lettre, montrent assez que la rupture avec les mathématiques habituelles est profondément sentie ⁽³⁾. « Je ne vois vraiment pas ce qui pourrait nous retenir dans cette activité créatrice de nouveaux nombres, aussitôt qu'il apparaît que pour le progrès de la science l'introduction d'une nouvelle, parmi ces innombrables classes de nombres, est devenue souhaitable ou même indispensable ; et la dernière éventualité me semble se produire effectivement en théorie des ensembles, peut-être même aussi encore dans un domaine beaucoup plus étendu, du moins *sans* cette extension je ne peux plus aller de l'avant, *avec* elle j'atteins toute sorte d'inattendu. »

(1) D'après une note manuscrite de DEDEKIND (brouillons de *Was sind und was sollen die Zahlen* ?) l'entrevue de HARZBURG eut lieu du 7 au 12 septembre (DEDEKIND communiqua la 1^{re} esquisse de son travail). La création de CANTOR fut donc très rapide, puisque le mémoire est daté d'octobre.

(2) CANTOR-DEDEKIND (I), p. 55.

(3) « Je ne me dissimule aucunement qu'avec cette entreprise je me mets dans une opposition certaine avec des vues répandues sur l'infini mathématique et avec des opinions souvent défendues sur l'essence de la grandeur numérique. » CANTOR (XV), p. 165.

De même le début du mémoire : sans cette extension du concept de nombre « il me serait impossible de faire le plus petit pas en avant dans la théorie des ensembles ; puisse cette circonstance me justifier, ou si c'est nécessaire, m'excuser d'introduire des idées en apparence étranges dans mes développements ⁽¹⁾ ».

Le surprenant n'est pas tant, en effet, dans l'usage de nouveaux symboles donnant lieu à un calcul spécial — comme toutes les extensions du nombre, et plus récemment les travaux de GRASSMANN et de KUMMER en avaient donné l'exemple ⁽²⁾ — ni même dans la considération de l'infini actuel, invoqué déjà en particulier pour compléter le plan d'une variable complexe, que dans la décomposition logique de la notion de nombre et la superposition de classes d'objets, hors des prises de toute atteinte effective. Définition du nombre comme caractère émanant d'un système posé en soi d'objets et se dissociant en deux concepts : « celui de la puissance... indépendante de l'ordre imposé à l'ensemble et celui du nombre ordinal (ANZAHL) qui est nécessairement lié à l'ordre imposé à l'ensemble par une règle, au moyen de laquelle l'ensemble devient... *bien ordonné*... Si je redescends de l'infini au fini je vois avec la même clarté et la même beauté comment les deux concepts redeviennent « un » et se *fondent* dans le concept du nombre entier fini ⁽³⁾. » La notion de *bon ordre* pour un ensemble, — être ordonné de telle façon qu'à tout sous-ensemble corresponde dans l'ensemble global un élément

(1) L'excuse est reprise vingt pages plus loin : « sans les concepts susdits [nombre ordinal et cardinal] on ne va — j'en ai la conviction — pas plus loin dans la théorie des ensembles ; et il en est de même, semble-t-il, des domaines qui dépendent de la théorie des ensembles ou ont avec elle le contact le plus intime comme, par exemple, la théorie moderne des fonctions d'une part, la logique et la théorie de la connaissance d'autre part ». *Ibid.*, p. 180-181.

(2) *Ibid.* p. 183.

(3) P. 181.

qui suit immédiatement — provient en effet de l'extension de l'opération de la numération telle que la provoque la dérivation : au premier principe d'engendrement des numéros d'ordre finis — adjonction d'une unité — doit être joint pour le numérotage des dérivés un second principe qui consiste, « quand on se trouve en présence d'une succession déterminée quelconque de nombres entiers réels définis parmi lesquels aucun n'est le plus grand, à créer un nouveau nombre qui sera considéré comme la *limite* de ces nombres, c'est-à-dire comme le nombre immédiatement plus grand qu'eux tous ⁽¹⁾ ». L'analogie avec la définition des nombres irrationnels ⁽²⁾ — reprise d'ailleurs dans le même mémoire — est évidente : seulement les nombres irrationnels sont justifiés en tant que nombres par le double caractère d'être insérables dans un ordre, où figurent d'ailleurs les représentants des nombres rationnels, et de permettre la définition entre eux des 4 opérations arithmétiques — obéissant aux lois ordinaires (et donnant les résultats ordinaires pour les représentants des nombres rationnels). Pour les nombres transfinis l'ordre est défini par leur processus d'engendrement : *le 2^e principe n'est autre que la condition du bon ordre*. Donc ensemble bien ordonné dénoté par un ordinal transfini et ensemble des ordinaux inférieurs à cet ordinal sont deux réalités coïncidant du point de vue de la théorie. Quant aux règles d'arithmétique, elles se déduisent immédiatement de la signification intuitive des opérations : ajouter un ensemble à un autre est former un nouvel ensemble où tous les éléments du second suivent tous les élé-

(1) CANTOR (XV), p. 196.

(2) L'analogie est d'ailleurs explicitement signalée dans la lettre à LASWITZ du 15 février 1884 (CANTOR (XXIII), p. 395) : « Les nombres transfinis sont dans un certain sens même de *nouvelles irrationalités* et en fait la meilleure méthode à mes yeux pour définir les nombres irrationnels finis est... je pourrais même dire dans le principe la même que ma méthode pour l'introduction des nombres transfinis. On peut dire absolument : les nombres transfinis *restent ou tombent* avec les irrationnels finis ».

ments du premier, l'ordre intérieur à chaque ensemble étant conservé. Multiplier un ensemble A par un ensemble B c'est remplacer tout élément de B par un ensemble équivalent à A et ordonné de même façon, l'ordre de B étant conservé. Sur les procédés de calcul, la décomposition en produit de nombres premiers transfinis, le mémoire ne donne encore que de brèves indications.

Mais le but à atteindre n'est pas une théorie de ces nombres : il s'agit des puissances dont le caractère numérique avait été posé dès l'abord. Or à la première difficulté, passage d'une des deux puissances connues à l'autre, s'en ajoute maintenant une seconde, comparabilité des puissances — qui avait été postulée sans plus en 1877. De la relation (1) d'inégalité définie alors —

puissance $M <$ puissance N si $M \sim N'$,

N' étant une partie de N et M non équivalent à N ... résulte que si M est applicable sur N' partie de N et N sur M' partie de M on devrait pouvoir démontrer que M et N ont même puissance ; ou encore en appelant M'' la partie de M' sur laquelle l'application de N sur M' applique N' :

$$M \sim N' \sim M'' \quad (1) ;$$

$$\text{si puissance } M'' \leq \text{puissance } M' \leq \text{puissance } M \quad (2)$$

on a :

$$\text{puissance } M' = \text{puissance } M = \text{puissance } M'' \quad (3).$$

C'est donc la suite (2) d'inégalités qu'il faudrait démontrer. « Vous vous rappelez, écrit CANTOR à DEDEKIND le 5 novembre (2), que je vous disais à Harzburg qu'il ne m'était pas possible de démontrer le théorème suivant : si M' est partie d'un ensemble M , M'' partie de M' et si... M et M'' ont même puissance, M' a aussi même puissance

(1) Nous nous servons du symbole $\sim$ pour exprimer l'équivalence de deux ensembles.

(2) CANTOR-DEDEKIND (I), p. 55.

de M et M'' . Maintenant j'ai trouvé la source de ce théorème. » Ce sont les nombres transfinis qui réduisent l'une et l'autre difficulté.

En premier lieu ils permettent d'engendrer régulièrement des puissances supérieures. On a vu que les symboles transfinis du précédent mémoire ne permettaient pas de dépasser le dénombrable ; les ordinaux nés des deux principes (que l'on appellera ordinaux de la classe II, la classe I étant réservée aux ordinaux finis), n'étant que leur nouveau déguisement ne dénoteront donc jamais que des ensembles dénombrables. Mais leur ensemble n'est pas dénombrable : en effet, à toute suite dénombrable d'ordinaux de la classe II correspond soit un ordinal maximum appartenant à la suite, soit un ordinal de la classe II immédiatement supérieur, c'est-à-dire tel que tout ordinal inférieur est dépassé par un terme de la suite ; si donc l'ensemble de tous les ordinaux pouvait se mettre sous forme de suite dénombrable, ou bien il y aurait un terme γ plus grand que tous les autres, alors $\gamma + 1$ n'appartiendrait pas à la suite, ou bien il y aurait un nombre β de la classe II immédiatement supérieur à la suite. On reconnaît dans ce raisonnement un nouveau procédé de dépassement des suites dénombrables : le premier, pour le continu, utilisait la propriété que toute suite convergente d'éléments définit un nouvel élément ; celui-ci que toute suite d'ordinaux de la classe II peut être bien ordonnée (donc définit un nouvel ordinal de la même classe), méthode qui donnera le paradoxe de BURALIFORTI (1). Donc est brisé le cercle magique qui enfermait les symboles transfinis. La classe II sera suivie d'autres classes à condition d'ajouter aux deux principes d'engendrement un *principe d'arrêt*, qui clôtüre les classes, donc donne une nouvelle application aux principes moteurs :

(1) V. ci-après, p. 94.

n'appartiennent à une classe que les ordinaux dénotant des ensembles ayant la puissance de la classe précédente ; ainsi, pour la classe II, des ensembles dénombrables. Donc le 2^e principe moteur définira un ordinal immédiatement supérieur à ceux de la classe II, Ω premier élément de la classe III. « Sur cette voie... on peut, avec la plus grande sécurité et évidence, parvenir à des classes de nombres toujours nouvelles, et avec elles à toutes les puissances successives différentes qui se présentent dans la nature corporelle et spirituelle (1). » En effet, l'échelle ainsi obtenue permet de passer d'une puissance à la puissance immédiatement supérieure ; le théorème est démontré pour la classe II : tout ensemble d'ordinaux de la classe II est ou fini, ou dénombrable ou a la puissance de la classe II ; *quantum non datur*. Donc le procédé est pleinement satisfaisant si on est sûr d'atteindre ainsi toutes les puissances.

C'est ici qu'apparaît la deuxième question. D'abord un postulat ou une promesse : « qu'il soit toujours possible de mettre tout ensemble *bien défini* sous forme d'un ensemble *bien ordonné*, sur cette loi, me semble-t-il, fondamentale et féconde de notre pensée, particulièrement remarquable par son universalité, je me propose de revenir dans un mémoire ultérieur (2) » — promesse non tenue. Puis une démonstration, restreinte au cas de la classe II. Le théorème de l'équivalence est immédiat si M a la puissance de la classe II : en effet, M' ne peut être que fini dénombrable ou de puissance de la classe II ; comme il possède une partie M'' ayant cette puissance, il lui est équivalent. Ici encore promesse de solution ultérieure pour le cas général : dans la lettre à DEDEKIND on a vu que CANTOR croyait déjà la tenir.

Mais, même si on admet « la loi fondamentale du bon ordre », le problème initial n'est pas résolu : dans l'échelle

(1) CANTOR (XV), p. 199.

(2) *Ibid.*, p. 169.

des puissances, quelle place a la puissance du continu ? Ici se produit, symétriquement à la création des ordinaux, un événement qui restreint au moins la recherche, l'apparition des ensembles parfaits dans le processus de dérivation. On a vu que tout ensemble, dont le 1^{er} dérivé est dénombrable, est lui-même dénombrable. Le mémoire d'octobre 1882 introduit un nouveau résultat, d'ailleurs inexact partiellement : si le dérivé $P^{(1)}$ n'est pas dénombrable il est décomposable d'une seule manière en un ensemble réductible R et un ensemble parfait S ; un ensemble parfait étant un ensemble identique à son premier dérivé. L'opération de la dérivation renouvelée un nombre suffisant de fois (nombre qui est un des ordinaux de la classe II) réduit tout ensemble dont le 1^{er} dérivé n'est pas dénombrable à un ensemble parfait. Le théorème était inexact, ainsi que le montra BENDIXSON ⁽¹⁾ au sujet de R qui peut être irréductible, mais il est toujours dénombrable et sans point commun avec ses dérivés à partir d'un certain rang, de sorte que l'essentiel du résultat subsiste. C'est le mémoire de novembre 1883 qui donne la démonstration et étudie la puissance des ensembles parfaits ⁽²⁾. D'abord un ensemble parfait ne peut être dénombrable : preuve le procédé des intervalles emboîtés ; le raisonnement est à peu près le même que pour le continu. On remplace la densité par la densité en soi (tout point de l'ensemble est point limite, propriété que possède évidemment le continu) ; les intervalles ne seront plus pris entre deux points, mais autour d'un point. En second lieu, tout ensemble dont le 1^{er} dérivé est dénombrable est lui-même réductible. Il suffit de se rappeler la décomposition opérée par la dérivation et que l'introduction des nombres ordinaux, avec Ω , rend plus simple :

(1) « Dans une lettre qu'il m'adressa (mai 1882). Sur ma demande il a publié les résultats trouvés à cette occasion dans les *Acta mathematica*, t. II, p. 415 ». CANTOR (XVIII), p. 224. Cf. BENDIXSON (I).

(2) CANTOR (XVIII).

$$P(1) \equiv \sum_{\gamma} [P(\gamma) - P(\gamma+1)] + P(\Omega) \quad (1)$$

(où γ parcourt tous les nombres des classes I et II). Si aucune des différences sous le Σ n'était nulle, le Σ aurait la puissance de la classe II ; donc elles s'annulent toutes à partir d'un certain γ ; d'autre part, on ne peut avoir :

$$P(\gamma) \equiv P(\gamma+1),$$

l'ensemble $P(\gamma)$ étant parfait et contenu dans $P(\gamma)$ celui-ci ne pourrait être dénombrable. Donc à partir d'un certain γ tous les dérivés $P(\gamma)$ sont $\equiv 0$. Enfin si le 1^{er} dérivé d'un ensemble a une puissance supérieure au dénombrable, il existe des points communs à tous les dérivés et leur ensemble est parfait. Le procédé utilisé pour la démonstration est une spécialisation du principe de la propriété distributive utilisée par BOLZANO dite ici propriété r (propriété telle que, si un ensemble la possède dans toute division finie de cet ensemble, une des parties au moins la possède et toute somme de cet ensemble avec un autre la possède également) : on démontre que dans tout ensemble M possédant une propriété r il existe au moins un point g dont tous les voisinages si petits soient-ils renferment des parties de M possédant r. Avoir tous ses dérivés différents de 0 est pour l'ensemble P une propriété r : le point g correspondant est donc point limite de tous les dérivés (puisqu'ils ont tous une portion contenue dans tous les voisinages de g) ; comme à partir de $P(1)$ ils contiennent tous leurs points limites, g leur appartient à tous. On voit d'autre part sans peine que l'ensemble des points tels que g est parfait. Il est donc possible de décomposer $P(1)$ en un ensemble dénombrable R et un ensemble parfait S :

$$P(1) \equiv R + S \quad (2)$$

La formule (1) dans laquelle $P(\Omega)$ est évidemment un ensemble parfait puisque commun à tous les dérivés, oblige

à identifier R et $\Sigma [P(\gamma) - P(\gamma+1)]$, donc γ ne parcourt qu'une infinité dénombrable de nombres de la classe II, il y a un plus petit ordinal α pour lequel :

$$P(\alpha) \equiv P(\alpha+1) \equiv S$$

Ainsi l'aboutissement à un ensemble parfait remplace le parcours effectif de la totalité de la classe II : rapport qui pouvait sembler de favorable augure pour l'hypothèse sur le continu. En effet un dernier résultat donnait au problème son acuité maximum : tout ensemble parfait a la puissance du continu. La démonstration s'effectue sur le cas particulier de l'ensemble parfait non dense (découvert par DU BOIS REYMOND) : les intervalles contigus à cet ensemble P forment dans l'intervalle global un ensemble dénombrable partout dense et chaque point de P peut être défini de façon univoque par une suite convergente de ces intervalles (il est limite de suites d'extrémités). Si donc on met en correspondance biunivoque les intervalles contigus à P et l'ensemble des nombres rationnels de l'intervalle $(0...1)$, à tout nombre réel dans $(0...1)$, défini par une suite de nombres rationnels, correspondra de façon univoque un point de P défini par la suite homologue d'intervalles — et réciproquement. D'autre part tout ensemble fermé — c'est-à-dire contenant ses points limites — peut être considéré comme le 1^{er} dérivé d'un ensemble : il est donc ou dénombrable ou de puissance du continu. Pour toute une catégorie d'ensembles est donc démontrée l'hypothèse de la dualité des puissances. Puisque la puissance de la classe II vient immédiatement après le dénombrable, si elle était identique à celle du continu, il n'y en aurait bien que deux. D'autre part si l'on pouvait construire effectivement un ensemble fermé ayant la puissance de la classe II, en vertu de l'égalité (2) le problème serait résolu. « Comme vous le voyez, écrit CANTOR à MITTAG-LEFFLER en 1884, tout revient à définir seulement un seul ensemble fermé

ayant la 2^e puissance (1). » C'est à quoi il s'acharne pendant toutes ces années, attaquant le problème surtout du biais de la théorie des ensembles de points, privilégiant parmi ceux-ci toujours les ensembles fermés, traversant d'ailleurs parfois des crises de doute qui lui font contester la possibilité d'égaliser continu et puissance d'une classe quelconque de nombres transfinis (2).

Dans cette période de regroupement et de systématisation, plutôt que de création, apparaît cependant un résultat nouveau sur les puissances (3) : la démonstration par le procédé de la diagonale que l'ensemble des fonctions d'une variable a une puissance supérieure à l'ensemble de variation de son argument. On a vu que DU BOIS REYMOND employait déjà de façon intuitive le procédé de la diagonale en 1876. CANTOR part du cas où l'ensemble de variation de l'argument est l'ensemble des entiers : le système des fonctions prenant deux valeurs sur cet ensemble est alors celui des nombres réels, nouvelle démonstration du théorème de 1874. Mais le procédé peut se généraliser à n'importe quel ensemble base ; c'est le troisième procédé de dépassement, plus efficace que les deux autres puisque non restreint à une suite dénombrable, comme le premier, ou à une suite bien ordonnée comme le deuxième, qui utilise l'indépendance relative des deux variables : l'ensemble des fonctions étant supposé coordonné à celui de l'argument, il est toujours possible de définir une nouvelle fonction par la propriété d'avoir, pour un argument quelconque, une valeur différente de celle prise par la fonction coordonnée à cet argument. CANTOR n'envisage comme

(1) L. du 26 août 1884, SCHÖENFLIES (II), p. 17.

(2) L. du 14 novembre à MITTAG LEFFLER : « Heureusement je n'ai utilisé nulle part cette affirmation [que le continu a la puissance de la classe II] comme moyen de démonstration... J'ai trouvé une démonstration rigoureuse que le continu n'a pas la puissance de la classe II et bien plus qu'elle n'a aucune puissance exprimable en général par un nombre. » Cf. CANTOR-DEDEKIND (I), p. 5.

(3) CANTOR (XXV).

autre application que le dépassement du continu : c'est l'ensemble de toutes les fonctions d'une variable réelle. Qu'il y ait dans le procédé exactement la réponse à la question posée en 1882 par les résultats mêmes : y a-t-il un moyen régulier pour engendrer des puissances successives, en particulier pour le passage entre les deux puissances connues (dont la mise en présence exigeait qu'il y ait détermination de leur rapport ?) donc inutilité de faire le détour des classes d'ordinaux transfinis, il ne semble pas le remarquer. Bien plus, il s'appuie, pour confirmer « que les puissances n'ont pas de maximum », sur la démonstration, obtenue « par de tout autres moyens » que « le système de toutes les puissances si nous les envisageons ordonnées suivant leur grandeur forme un ensemble bien ordonné de sorte qu'à toute puissance correspond dans la nature une puissance immédiatement supérieure, mais aussi à toute suite croissante illimitée de puissances une qui la dépasse immédiatement ⁽¹⁾ ». Tant étaient forts à la fois l'attrait de l'opération du bon ordre et la confiance dans l'hypothèse du continu. Aucun des trois principes présidant à l'engendrement des classes d'ordinaux transfinis ne permet d'ailleurs de prouver la dernière affirmation, c'est-à-dire d'atteindre seulement la classe d'ordinaux qui a pour numéro d'ordre ω . C'est l'obsession imaginative du processus engendré par l'opération de la dérivation : le lien entre les développements de la théorie qui en naissent et le problème effectivement posé par la découverte des deux puissances apparaît donc cette fois contingent ; il s'agit de deux manières radicalement différentes d'envisager les ensembles, aucun problème effectivement posé n'assure qu'il y ait communication de l'une à l'autre.

(1) *Ibid.*, p. 280.

* * *

De là une dualité latente, un malaise sous-jacent dans la systématisation de la science de l'infini à laquelle est consacré le reste des travaux cantorien. Il s'agit, comme le note ZERMELO ⁽¹⁾, uniquement « de développement » et de précision d'idées dont la « quintessence » se trouve dans les six mémoires de 1883-1884. D'une part théorie de l'infini actuel comme fondement de l'édifice, d'autre part les deux arithmétiques cardinale et ordinale du transfini, avec, s'ajoutant à la seconde, la nouvelle théorie des types d'ordre. La caractéristique de ces derniers exposés est d'achever l'identification entre puissances, nombres transfinis d'une part, nombres de l'arithmétique ordinaire d'autre part : les règles de calcul simplement indiquées en 1883 sont développées et précisées, le rapport aux collections d'objets est posé comme le même que pour les nombres finis.

Théorie réalisée de l'infini actuel qui se poursuit à travers la correspondance rassemblée dans les *Communications sur la doctrine du transfini*. Comme le fait observer FRAENKEL, à aucun moment il ne s'agit d'une construction arbitraire de théorie abstraite justifiée par la seule non contradiction de ses axiomes. CANTOR avait bien fait, en 1883, la distinction entre *réalité immanente* ou *intrasubjective* et *réalité transiente* ou *transsubjective*, la seconde réservée aux objets de la physique, tandis que les objets de la mathématique « dont l'essence est la liberté » ⁽²⁾ ne peuvent prétendre qu'à la première, puisque leur création n'est liée que par « la restriction, qui va de soi, que leurs concepts soient non contradictoires ». Mais dans le même passage apparaît aussi un nouveau critère pour qu'à « l'arbitraire soit laissé le moindre jeu..., chaque concept mathématique porte en

(1) CANTOR (XXVIII), p. 246.

(2) CANTOR (XV), p. 182.

lui le correctif nécessaire ; est-il infructueux ou incommodé, il le manifeste bientôt par son caractère inutilisable (1) ». Cette fécondité ou stérilité possible n'a-t-elle pas son fondement dans la nature, même la nature physique ? A côté de ses applications à la mécanique, à la physique, à la biologie — sur lesquelles CANTOR revient à plusieurs reprises — la justification de la théorie est dans le fait de l'existence d'ensembles infinis, comme les points rationnels ou irrationnels d'une droite. Des définitions de la puissance, puis des nombres ordinaux et cardinaux le montrent assez : « j'entends par *puissance* ou nombre cardinal d'un ensemble M , écrit-il à LASSWITZ en 1884, le concept universel ou générique (universale) que l'on obtient en faisant abstraction pour l'ensemble aussi bien de la constitution de ses éléments que de toutes les relations que ces éléments ont entre eux ou avec d'autres choses donc en particulier aussi de l'ordre qui règne entre eux, et ne considère que ce qui est commun à tous les ensembles *équivalents* à M (2). » Il ne s'agit même pas d'une définition par postulats, relative à une opération — l'application — comme semble l'indiquer la dernière phrase. Ailleurs — même mémoire — la puissance de l'ensemble est le « concept général dans lequel les éléments comme des sortes d'*uns* sont réunis en un tout unifié, de sorte que chacun d'eux n'a pas de rang privilégié (3) ». Définition reprise en 1897 — contre laquelle pouvait s'élever ZERMELO (4) parce qu'impropre à fonder une science abstraite — mais caractéristique de la théorie en tant que réaliste : « le nombre cardinal est lui-même un ensemble déterminé formé de la réunion de simples *uns* qui a son existence dans notre esprit comme image intellectuelle ou pro-

(1) *Ibid.*

(2) Lettre publiée in CANTOR (XXIII), p. 387.

(3) *Ibid.*, p. 379.

(4) CANTOR (XXVIII), p. 351.

jection de l'ensemble M ⁽¹⁾ ». Si les *uns* forment une pluralité quoique rien ne les distingue plus entre eux, c'est qu'ils ne sont que la silhouette incolore des objets concrets eux-mêmes pour lesquels on tient compte du fait qu'ils sont distincts, en négligeant le contenu de cette distinction. D'où l'importance de la question qu'il y ait effectivement un infini actuel, la recherche des cautions dans l'histoire, depuis PYTHAGORE. D'où également le rejet de l'infiniment petit actuel dont STOLZ ⁽²⁾, DU BOIS REYMOND ⁽³⁾ et VÉRONÈSE ⁽⁴⁾ essayaient de faire la théorie. Non pas seulement les difficultés techniques de celle-ci, l'impossibilité, par exemple, d'utiliser la division par nombres transfinis puisqu'on n'obtient pas de résultat univoque, mais surtout l'origine intuitive de la théorie des ensembles et l'exigence de ces méthodes protestent contre toute assimilation ⁽⁵⁾. Elle part, note FRAENKEL, « du concept dans une certaine mesure intuitif d'ensemble... la mise en ordre [des nombres transfinis] et les lois du calcul sont par suite non pas décidées par des actes arbitraires de définition, mais dans une certaine mesure décelées dans la nature, avec leur structure imposée. Ainsi apparaissent les règles de calcul très hétérogènes entre elles pour diverses espèces de nombres qu'il n'y aurait eu, *a priori*, aucune raison de poser ainsi et... [auxquelles]... par des considérations philosophiques sur l'infiniment grand on ne serait certainement pas venu ⁽⁶⁾ ». Parmi ces lois, l'une des plus importantes est celle qui pose une limite à toute suite de nombres croissants, loi qui est une sorte d'extension de l'acte de compter :

(1) CANTOR (XXVI), p. 283.

(2) STOLZ (II).

(3) DU BOIS REYMOND (IX).

(4) VÉRONÈSE (I).

(5) « Le fait de l'existence des nombres infiniment grands est donc si peu une raison pour l'existence des grandeurs infiniment petites que bien plus justement avec l'aide de ceux-là l'impossibilité de ceux-ci est démontrée. » CANTOR (XXIII), p. 408.

(6) FRAENKEL (X), p. 118-119.

définir après toute suite infinie un nombre immédiatement supérieur, affirmer en quelque sorte l'empire sans borne des nombres, axiome d'ARCHIMÈDE prolongé. D'où la seconde raison pour rejeter l'infiniment petit : il ne s'insère que dans un système non archimédique. « Ce qu'on appelle l'*axiome d'Archimède* n'est pas un axiome, mais une proposition découlant avec une nécessité logique du concept de grandeur linéaire ⁽¹⁾. »

Les lois naturelles du calcul avec les nombre cardinaux supposent donc au préalable une définition des rapports des ensembles entre eux lorsqu'on fait abstraction de l'ordre : ainsi sont introduites en 1895 les notions de partie d'ensemble — restreinte au cas de la vraie partie (c'est-à-dire distincte du tout) — de réunion, ou ensemble somme (restreinte à la réunion d'ensembles sans éléments communs), d'ensemble produit de deux ensembles M et N : ensemble de tous les couples d'éléments pris l'un dans M l'autre dans N (définition évidemment équivalente à celle qui avait été donnée en 1883) ; enfin de l'exponentiation d'un ensemble par un autre : ensemble de toutes les *façons de recouvrir* (BELEGUNGEN) un ensemble par les éléments d'un autre, c'est-à-dire de toutes les fonctions dont l'argument est pris dans le premier ensemble et la valeur dans le second. Opérations concrètes, que complète l'acte également concret de la coordination biunivoque des éléments de deux ensembles équivalents. En regard, les rapports qui s'en déduisent pour les nombres cardinaux de ces ensembles : de l'équivalence se déduit l'égalité, ce qui doit être démontré puisque le processus d'abstraction définissant le nombre est posé comme différent du processus d'application (différence imaginative, qu'il est caractéristique de voir CANTOR maintenir ici). De la relation entre ensemble et partie, les inégalités numériques : « s'il n'y a aucune partie de M équiva-

(1) CANTOR (XXIII), p. 409.

lente à N et s'il y a une partie N_1 de N équivalente à M », on pose ⁽¹⁾ :

$$\overline{\overline{M}} < \overline{N}$$

transitivité, exclusion mutuelle des 3 relations $=$, $>$, $<$ sont démontrées. En revanche « il ne va nullement de soi et ne pouvait être ici démontré que pour tout couple de nombres cardinaux une de ces 3 relations doive être nécessairement réalisée ⁽²⁾ ». C'est la première fois que la comparabilité universelle ne fait pas l'objet d'une simple affirmation ; les corollaires, dont l'établissement est réservé avec elle, car « ils s'en déduisent immédiatement », précisent la distinction |des deux autres cas qui peuvent se présenter d'un point de vue purement logique, dans les relations entre deux ensembles :

« il y a une partie N_1 de N équivalente à M et il y a aussi une partie M_1 de M équivalente à N ;

il n'y a ni une partie de N équivalente à M ni une partie de M équivalente à N ⁽³⁾. »

L'un des corollaires exclut la deuxième hypothèse ; l'autre affirme que dans la première $\overline{\overline{M}} = \overline{N}$: c'est le théorème de l'équivalence dont SCHRÖDER et BERNSTEIN devaient donner la démonstration sans invoquer la comparabilité ⁽⁴⁾.

(1) CANTOR représente la puissance d'un ensemble M par $\overline{\overline{M}}$ pour rappeler la double abstraction (propriétés des éléments et leur ordre) et son type d'ordre par $\overline{M}$ — puisqu'on ne fait plus qu'une abstraction.

(2) CANTOR (XXVI), p. 285.

(3) L. à DEDEKIND, 30 août 1899, CANTOR (XXVIII), p. 450.

(4) La démonstration est d'ordre élémentaire (cf. l'exposé qu'en donne, en 1898, E. BOREL (I), p. 104-106). Soient M , T et P :

$$P \infty M \text{ et } P \subset T \subset M.$$

$$\text{Posons } M - T = R \quad T - P = Q$$

L'application f de M sur P donne :

$$\begin{aligned} P &= f(P + Q + R) = f(P) + f(Q) + f(R) = P_1 + Q_1 + R_1 \\ P_1 &= f(P_1 + Q_1 + R_1) = f(P_1) + f(Q_1) + f(R_1) = P_2 + Q_2 + R_2 \text{ etc.} \end{aligned}$$

en posant $P_1 = f(P)$, etc.

Si on pose P_ω = intersection de $P, P_1, P_2, \dots$ on a :

$$\begin{aligned} P &= P_\omega + Q_1 + R_1 + Q_2 + R_2 + Q_3 + R_3 + \dots \infty \\ &\infty P_\omega + Q + R_1 + Q_1 + R_2 + Q_2 + R_3 + \dots = Q + P = T. \end{aligned}$$

puisque $P_{n-1} - P_n = Q_n + R_n$

(HAUSDORFF (I), p. 50).

Pour celle-ci, déclare CANTOR, « sa vérité ne pourra être reconnue que plus tard lorsque nous aurons un aperçu de la suite croissante des nombres cardinaux transfinis et une idée de leur enchaînement », c'est-à-dire après l'exposé de la théorie des nombres ordinaux. Difficulté dans le plan qui en manifeste une plus grave : le développement, — et la justification interne de la théorie telle qu'elle est présentée exige que l'on ne fasse appel qu'aux intuitions évoquées par les définitions des rapports entre ensembles. Utiliser la théorie du bon ordre est, sinon commettre un cercle vicieux — puisque les résultats invoqués dans la théorie ordinaire sont préalables à l'affirmation de la comparabilité — du moins porter atteinte aux fondements de l'édifice : il ne subsiste que si les intuitions naïves originelles lui suffisent. D'où le recours ultérieur à l'axiomatisation : c'est la difficulté de la comparabilité, avant les paradoxes, qui en est la première raison.

Les résultats effectifs de la théorie cardinale ne peuvent être qu'assez pauvres : on se borne à retrouver les propriétés habituelles de l'addition et de la multiplication (commutativité, associativité, distributivité), de l'exponentiation :

$$a^b \cdot a^c = a^{b+c} \quad (a^b)^c = a^{bc}$$

Pour aller plus loin la première distinction qui s'impose est celle des opérations effectuées sur un nombre fini ou un nombre infini de termes, c'est-à-dire, puisque la théorie prétend être préalable à toute discipline mathématique, l'étude des ensembles finis et celle de l'infini dénombrable.

La théorie des ensembles finis manque de rigueur. Dans les *Communications sur la doctrine du transfini* ⁽¹⁾ (1884), la définition donnée suppose déjà une notion d'ordre : « nous entendons par ensemble fini un ensemble M qui est engendré par une telle adjonction successive d'éléments à

(1) CANTOR (X.XIII) ; le mémoire a été publié en 1887 mais est daté de 1884.

un élément originairement donné, que *inversement* par éloignement successif des éléments dans l'ordre opposé, l'élément origine peut être retrouvé ⁽¹⁾ ». Cette évidente difficulté est évitée par suppression de toute définition générale : les nombres cardinaux finis sont obtenus chacun à partir du précédent par adjonction d'un objet à l'ensemble qu'il dénotait. On empêche par artifice que, comme indice du nouvel objet, figure le nombre cardinal dans sa propre définition, mais on ne peut éviter que l'emploi de l'induction complète pour les propriétés élémentaires — comme : toute partie d'un ensemble ayant pour nombre cardinal ν possède pour nombre cardinal un de ceux qui précèdent ν — représente un cercle vicieux. La seule démonstration effective est celle de la proposition que si un ensemble M n'est équivalent à aucune de ses parties, l'ensemble $\{M, e\}$ formé par adjonction à M d'un nouvel élément e possède la même propriété ⁽²⁾. En 1884 la propriété est signalée comme « caractéristique des ensembles finis », donc pouvant servir de définition, ce que faisait déjà DEDEKIND. Mais la notion d'engendrement est essentielle chez CANTOR. Le plus petit infini est construit : « l'ensemble de tous les nombres cardinaux finis nous offre l'exemple le plus proche d'un ensemble transfini ». Il est celui dont le nombre cardinal est le plus petit ; en effet tout ensemble transfini possède un sous-ensemble qui lui est équivalent : il suffit d'observer que la suppression dans un ensemble transfini d'un nombre fini d'éléments laisse toujours la possibilité d'en supprimer un autre ; l'ensemble de toutes les suppressions est donc l'ensemble des nombres cardinaux finis, c'est-à-dire le dénombrable. La propriété caractéristique du transfini est intuitive pour le dénombrable : on peut supprimer un élément de

(1) *Ibid.*, p. 415n.

(2) Le raisonnement employé est exactement celui qui sert à DEDEKIND in DEDEKIND (III).

l'ensemble sans que sa puissance change, il est donc équivalent à une de ses parties. Propriété qu'il transmet à tous les ensembles transfinis puisqu'il en est sous-ensemble. Enfin si on désigne par $\aleph_0$ son nombre cardinal on a les équations :

$$\aleph_0 + n = \aleph_0 \quad \aleph_0 + \aleph_0 = \aleph_0 \quad \aleph_0 \cdot n = \aleph_0 \quad \aleph_0 \cdot \aleph_0 = \aleph_0$$

Le continu, puisque ensemble, par exemple, de tous les développements dyadiques illimités, c'est-à-dire de toutes les fonctions dont l'argument parcourt un ensemble dénombrable et qui prennent les valeurs 0 et 1, a pour nombre cardinal $2^{\aleph_0}$. D'où en particulier une nouvelle démonstration de l'invariance de la puissance des continus quel que soit leur nombre de dimensions (fini ou infini dénombrable) :

$$(2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0}$$

L'étude n'est pas poussée plus loin ; aussitôt se pose la question du nombre cardinal suivant immédiatement $\aleph_0$: renvoi à la notion d'ordre.

La première mention des types d'ordre est faite dans le mémoire *Sur différents points de vue relatifs à l'infini actuel* : le point de vue abstrait nous permet de saisir l'infini actuel « sous forme... des nombres transfinis ou sous la forme encore plus générale des *types d'ordre transfinis* » (1). A la fin des *Communications sur la doctrine du transfini* se trouve esquissée une étude générale de ces types, dont les nombres ordinaux ne sont qu'une espèce particulière : le type est « le concept universel qui se dégage d'un ensemble M — si nous faisons abstraction seulement de la constitution des éléments, mais conservons les relations d'ordre entre eux ». L'élargissement de la théorie ordinale (qui ne considérerait que les ensembles bien ordonnés) en théorie des

(1) CANTOR (XXII), p. 372.

types est nécessaire pour une caractérisation abstraite du continu : la dissociation entre topologie et arithmétique transfinie s'achève entre 1884 et 1887. S'il est possible d'édifier de façon autonome une théorie des puissances où s'insère le continu, il faut que celui-ci puisse être, en tant que nombre, défini indépendamment de toute considération géométrique ; aussi bien l'opération simple de la dérivation s'était-elle montrée insuffisante : la solution sera cherchée dans l'analyse, qu'elle amorçait, de l'ordre entre éléments. Le mémoire de 1887 ne s'occupe que de distinguer entre ordres multiples, contre-coup des efforts pour définir la notion de dimension. En 1895 CANTOR revient au seul ordre linéaire et démontre les deux importants théorèmes qui caractérisent les types de l'ensemble des nombres rationnels et du continu.

Au préalable, une théorie générale des types : l'opération correspondant à l'application dans la théorie cardinale et définissant l'égalité des types est l'application *semblable*, conservant les relations d'ordre entre éléments. Puisqu'elle est cas particulier de la 1^{re}, on voit qu'à tout nombre cardinal $\aleph$ correspond la classe $[\aleph]$ de tous les types qu'une application non semblable transforme les uns dans les autres. De là deux nouvelles notions : celle de *type inverse*, ou conjugué — déjà introduite en 1887 — obtenu en inversant toutes les relations d'ordre dans un ensemble : ainsi l'inverse du type ω de l'ensemble des nombres entiers rangés par ordre de grandeur est $^*\omega$ type de l'ensemble des nombres entiers négatifs rangés par ordre de grandeur. Celle d'*automorphisme* : application semblable d'un ensemble sur lui-même distincte de l'identité ; certains types ω (en général tous les types d'ensembles bien ordonnés) ne permettent pas l'automorphisme, au contraire celui de l'ensemble de tous les entiers (négatifs et positifs), des nombres rationnels, etc. (en général ceux qui ne définissent ni premier ni dernier élément) sont applicables d'une infi-

nité de manières sur eux-mêmes. L'addition et la multiplication se définissent comme pour les ordinaux en 1883 ; pour le produit nouvelle présentation : si M est le multiplicande, N le multiplicateur, on considère l'ensemble des éléments m_n parcourant pour chaque indice n un ensemble semblable à M , l'indice parcourant N et l'ordre introduit par lui l'emportant sur l'ordre entre 2 éléments m et m' de M ; on peut ainsi généraliser au produit de k facteurs, l'élément est de la forme $m_{n_1 n_2 \dots n_k}$. Comme on voit immédiate-

ment, l'addition n'est pas commutative non plus que la multiplication ; toutes deux sont associatives et la multiplication possède la distributivité du multiplicateur :

$$\overline{M} \cdot (\overline{N} + \overline{P}) = \overline{M} \cdot \overline{N} + \overline{M} \cdot \overline{P}$$

Le type η de l'ensemble des nombres rationnels rangés par ordre de grandeur est défini par les 3 propriétés : 1^o appartenir à la classe $[\aleph_0]$; 2^o n'avoir ni premier ni dernier élément ; 3^o être tel qu'entre 2 éléments de l'ensemble l'ordre en insère d'autres : *densité*. La 3^e propriété est la première transposition abstraite d'une propriété topologique. La démonstration que tout ensemble possédant ces trois propriétés est *semblable* à l'ensemble des nombres rationnels s'effectue grâce au procédé qui avait procuré l'application du continu sur tout ensemble parfait non dense : la première étape réalisait une application semblable de l'ensemble des nombres rationnels sur celui des intervalles contigus à l'ensemble parfait non dense. Ici ce dernier ensemble est remplacé par un ensemble quelconque M possédant les trois propriétés de la définition : la 3^e propriété, densité, permet de coordonner progressivement (la possibilité d'une coordination progressive venant de la 1^{re} propriété : les deux ensembles peuvent être rangés en une suite numérotée par les entiers positifs), en respectant l'ordre, à tout nombre rationnel un élément de M et un seul, et réciproquement

de prouver que tout élément de M est ainsi pris dans la coordination.

Pour atteindre le continu, CANTOR décalque de même dans l'abstrait ses raisonnements sur les nombres réels : aux suites convergentes fondamentales définissant un nombre irrationnel correspondent ici des suites fondamentales abstraites croissantes ou décroissantes : l'ordre d'un ensemble étant conservé entre les éléments des ensembles partiels distingués en son sein, ceux-ci possèdent par là même un type d'ordre déterminé (en général distinct du type de l'ensemble total) ; on appellera suites fondamentales croissantes les ensembles partiels de type ω , suites fondamentales décroissantes les ensembles de type $^*\omega$. Deux suites, croissantes ou décroissantes toutes deux, sont conjuguées si à tout élément de l'une correspond dans l'autre un élément supérieur dans l'ordre. Lorsque l'une est croissante et l'autre décroissante, elles sont conjuguées : si tout élément de la 1^{re} précède tout élément de la 2^e et si entre elles l'ordre insère au maximum un élément de l'ensemble total M étranger à chacune d'elles. Une suite fondamentale possède dans M un élément limite -- ou élément principal dans M -- s'il existe un élément m_0 tel que pour tout élément a_v de la suite on ait :

$a_v \prec m_0$ (suite croissante) ou $a_v \succ m_0$ (suite décroissante) et que, si pour tout élément $m \prec m_0$ dans le 1^{er} cas, $m \succ m_0$ dans le 2^e, on puisse trouver dans la suite un rang v_0 tel que pour tout $v \geq v_0$ on ait :

$$m \prec a_v \text{ dans le 1^{er} cas} \quad m \succ a_v \text{ dans le 2^e}. \quad \text{---}$$

On voit qu'une suite fondamentale — avec toutes les suites qui lui sont conjuguées — ne peut posséder qu'un seul élément limite et que dans l'application semblable les caractères suite fondamentale et élément limite sont invariants. Enfin les notions de densité en soi, fermeture, perfection subissent la même transformation abstraite : un ensemble

ordonné est *dense en soi* si tous ses éléments sont éléments principaux, *fermé* s'il contient un élément limite pour toute suite fondamentale, *parfait* s'il est à la fois dense en soi et fermé. D'où la caractérisation du *type* θ du continu linéaire : c'est un ensemble ordonné M : 1^o parfait ; 2^o contenant un ensemble S dénombrable dense dans M . En supprimant de S le premier et le dernier élément s'ils existent, ce qui ne change pas la propriété de S utilisée dans l'énoncé, on voit que S a le type η . La démonstration reproduit la 2^e partie de la démonstration appliquant le continu sur un ensemble parfait non dense. L'application semblable de S sur l'ensemble R des nombres rationnels étant réalisée d'après le 1^{er} théorème, on détermine la correspondance entre les autres éléments par celle déjà établie entre les suites fondamentales d'éléments de R et de S dont ils sont respectivement éléments limites. Comme le remarque ZERMELO ⁽¹⁾, la démonstration est plus simple avec la définition DEDEKIND du continu comme ensemble ordonné sans saut (c'est-à-dire dense) ni lacune (c'est-à-dire tel que toute coupure soit déterminée par un élément). D'autre part la 1^{re} propriété exigée de θ est trop forte ; il suffit de le définir comme fermé puisque d'après la 2^e propriété il est dense et que la densité entraîne évidemment la densité en soi ⁽²⁾.

Le premier objectif de la théorie du bon ordre est l'établissement de l'universelle comparabilité des ensembles bien ordonnés. Il est atteint grâce à la notion de *segment* : ensemble des éléments précédant dans l'ordre un élément d'un ensemble ; le complémentaire du segment est son *reste*. Le caractère unique d'un segment dans un ensemble — impossibilité d'établir une application semblable entre

(1) CANTOR (XXVIII), p. 354.

(2) Si entre deux éléments d'un ensemble il y en a toujours d'autres, tout point sera point limite.

lui et, soit l'ensemble tout entier, soit un autre segment — est démontré un peu longuement par CANTOR : ZERMELO simplifia depuis, grâce au théorème que, dans toute application semblable d'un ensemble bien ordonné sur une de ses parties, aucun élément ne peut être appliqué sur un élément qui le précède ⁽¹⁾. Comme la similitude est transitive, il n'y a qu'un moyen d'appliquer semblablement un ensemble M sur un autre N ou sur un de ses segments N . La mise en correspondance des segments d'un ensemble M avec ceux d'un ensemble N — par application semblable des segments homologues l'un sur l'autre — définit une application semblable de M sur N et réciproquement.

Enfin le résultat cherché ; étant donnés deux ensembles bien ordonnés M et N , trois cas seulement sont possibles : ou M est semblable à N , ou M est semblable à un segment de N , ou N à un segment de M . Qu'ils soient chacun semblables à un segment de l'autre est impossible en vertu des théorèmes précédents. Qu'ils ne soient ni semblables entre eux, ni l'un semblable à un segment de l'autre est également exclu : en effet si tous les segments de M ont des homologues semblables dans N et réciproquement on a vu que la similitude de M et N s'en déduirait ; si un segment de M n'a pas d'homologue semblable dans N il y en a un plus petit M_1 , jouissant de cette propriété : ou bien tout segment de N a un segment semblable dans M (donc dans M_1) donc M_1 est semblable à N ; ou bien il y a également des segments de N sans homologue semblable dans M , soit

(1) En effet, le sous-ensemble des éléments jouissant de cette propriété posséderait un premier élément a , donc on aurait (f étant la fonction application) :

$$f(a) = a' \prec a ;$$

donc comme l'application est semblable $f(a') \prec f(a) = a'$ donc l'élément a' précédant a jouirait de la propriété, contrairement à la définition de a . On voit qu'une application semblable d'un segment sur un autre segment du même ensemble produirait ce résultat.

N_1 le plus petit : M_1 et N_1 sont semblables (puisque par hypothèse tous les segments plus petits, donc tous leurs segments, s'appliquent semblablement les uns sur les autres) contrairement à leur définition.

Par la comparabilité des ensembles, est en même temps définie celle des types d'ordre ou nombres ordinaux, qui les dénotent : une et une seule des trois relations ⁽¹⁾ $=, >, <$ existe nécessairement entre deux d'entre eux. L'arithmétique transfinie qui en découle procure les trois derniers objectifs du mémoire : démonstration que tout ensemble de nombres ordinaux rangés par ordre de grandeur est bien ordonné ; règles précises de calcul et de décomposition canonique d'un nombre en un produit de facteurs irréductibles ; enfin exploration de la classe II avec l'un des résultats les plus intéressants, définition des *nombres ε* . Mais dès ce moment se manifeste la gêne introduite par la découverte du paradoxe de BURALI-FORTI : à aucun des raisonnements faits ici n'est accordée sa portée véritable ; sauf la définition générale des opérations, qui reproduit les définitions données pour les types d'ordre, avec en plus chaque fois la démonstration que l'ensemble résultat est également bien ordonné (ce qui est évident aussi bien pour le produit que pour la somme), toutes les considérations sont restreintes aux nombres ordinaux de la classe II, alors que la plupart seraient immédiatement valables pour les ordinaux de n'importe quelle classe. Ainsi l'ensemble de tous les ordinaux inférieurs à un nombre ordinal donné α est bien ordonné et de type α : soit M un ensemble bien ordonné de type α , à tout nombre $\alpha' < \alpha$ correspond un segment de M et réciproquement, les éléments déterminent les segments de M , et les nombres inférieurs à α peuvent donc être mis en correspondance biunivoque, application qui conserve évidemment l'ordre. Si on considère

(1) Nous remplacerons désormais la relation $\prec$ par $<$ puisque le nombre ordinal a tous les caractères d'un nombre.

un ensemble quelconque P de nombres ordinaux transfinis (en particulier la classe II), l'intersection de chacune de ses parties (et de P lui-même) avec le segment déterminé par un élément de P possède un 1^{er} élément qui est évidemment aussi 1^{er} élément de la partie (ou de P) donc P est bien ordonné.

Les opérations n'avaient en principe été définies que pour un nombre fini de termes : on voit sans peine que la réunion d'une suite infinie d'ensembles bien ordonnés est bien ordonnée ; de là, en particulier, les notions de suite fondamentale (c'est-à-dire croissante) de nombres ordinaux et de limite correspondante : soit une suite $\alpha_1, \alpha_2, \dots$; si on fait la somme :

$$\beta = \alpha_1 + (\alpha_2 - \alpha_1) + (\alpha_3 - \alpha_2) + \dots$$

β est un nombre ordinal supérieur immédiatement à la suite, c'est-à-dire tel que tout $\alpha_v < \beta$ et que pour tout nombre $\gamma < \beta$ il y a des $\alpha_v > \gamma$, on posera :

$$\beta = \lim_v \alpha_v.$$

Tout nombre de la classe II est obtenu soit par adjonction de 1 à un nombre déjà connu (il est dit de 1^{re} espèce), soit par passage à la limite d'une suite fondamentale (2^e espèce) : de chaque façon on a donc le nombre qui suit immédiatement les nombres déjà posés (on reconnaît les deux principes d'engendrement du mémoire de 1883).

L'exponentiation n'avait été définie que pour le cas d'exposants finis, au delà l'intuition ne peut servir. C'est alors qu'intervient pour la première fois la récurrence transfinie. « Il y a une et une seule fonction complètement déterminée $f(\xi)$ de la variable ξ [parcourant la classe II] qui satisfait aux conditions suivantes (γ et δ étant des constantes, $\delta > 0$, $\gamma > 1$) :

$$1) f(0) = \delta$$

$$2) \text{ si } \xi' < \xi'' \quad f(\xi') < f(\xi'')$$

3) pour tout ξ $f(\xi + 1) = f(\xi) \cdot \gamma$

4) si $\{\xi_v\}$ est une suite fondamentale quelconque, $\{f(\xi_v)\}$ est aussi une suite fondamentale et si on a :

$$\lim_v \xi_v = \xi \quad \text{on a} \quad \lim_v f(\xi_v) = f(\xi). \text{ »}$$

La démonstration se fait par induction transfinie : pour les ξ finis ($\xi = 1, 2, 3, \dots$) on obtient de proche en proche $f(\xi)$ grâce à la condition 3) ; et la condition 2) procure pour $f(\omega)$ et de façon générale pour tout nombre de 2^e espèce des suites fondamentales qui déterminent f de façon univoque. Le bon ordre garantit l'extension du procédé à toute la classe II : si des nombres y échappaient, il y en aurait un plus petit α , mais les deux procédés 3) et 4) permettraient alors de calculer $f(\alpha)$. On voit la dépendance de la nouvelle induction par rapport aux deux principes d'engendrement (qui impliquent la notion de bon ordre) : toute propriété qu'ils laissent invariante sera démontrable par elle.

Si on pose $\delta = 1$ on écrit :

$$f(\xi) = \gamma^\xi;$$

en effet dans le cas de ξ fini on retrouve bien la définition ordinaire ; d'autre part les propriétés caractéristiques sont conservées :

$$\gamma^\eta \cdot \gamma^\xi = \gamma^{\eta + \xi} \quad (\gamma^\eta)^\xi = \gamma^{\eta \cdot \xi};$$

pour démontrer la 1^{re} égalité il suffit en effet de poser $\delta = \gamma^\eta$ et de considérer la fonction $f(\xi) = \gamma^{\eta + \xi}$: on voit que $\gamma^\eta \cdot \gamma^\xi$ satisfait aux quatre conditions ; on a donc l'égalité en vertu de l'unicité ; de même pour la 2^e, si on pose $\delta = 1$ et si on remplace γ par γ^η . D'une façon générale toute fonction satisfaisant aux quatre conditions est de la forme :

$$f(\xi) = \delta \cdot \gamma^\xi.$$

Les règles du calcul se caractérisent par leur harmonieuse simplicité, qui découle de la distributivité du multiplica-

teur dans la multiplication et de la propriété de la somme.

$$\alpha + \beta = \beta \quad \text{si } \alpha < \beta$$

on en déduit d'abord la forme normale univoque sous laquelle peut se mettre tout nombre de la classe II.

$$\alpha = \omega^{\alpha_0} k_0 + \omega^{\alpha_1} k_1 + \dots + \omega^{\alpha_r} k_r \quad (\text{I})$$

où les $k_0, k_1 \dots k_r$ sont des entiers finis, les $\alpha_0 \dots \alpha_r$ des nombres de la 1^{re} ou de la 2^e classe tels que :

$$\alpha \geq \alpha_0 > \alpha_1 > \dots > \alpha_r \geq 0$$

Il suffit de procéder de proche en proche : parmi les nombres ξ tels que $\omega^\xi > \alpha$, il y en a un plus petit β , qui est de 1^{re} espèce, on pose $\beta = \alpha_0 + 1$; on a donc :

$$\omega^{\alpha_0} \leq \alpha < \omega^{\alpha_0 + 1}$$

et la 2^e inégalité montre qu'il y a un plus grand nombre fini k_0 tel que :

$$\omega^{\alpha_0} k_0 \leq \alpha$$

on obtient donc :

$$\alpha = \omega^{\alpha_0} k_0 + \alpha' \quad \text{avec} \quad \alpha' < \omega^{\alpha_0} < \alpha$$

et ainsi de suite : l'opération ne peut se renouveler qu'un nombre fini de fois puisque la suite $\alpha > \alpha' > \dots$ est décroissante.

α_0 est appelé le degré de α . D'où la définition des quatre opérations effectuées sur les formes normales ; dans l'addition ne sont conservés du 1^{er} sommant que les termes de degré supérieur au degré du 2^e :

$$\alpha + \beta = \omega^{\alpha_0} k_0 + \dots + \omega^{\alpha_p} k_p + \omega^{\beta_0} \lambda_0 + \dots + \omega^{\beta_\sigma} \lambda_\sigma.$$

Dans la multiplication si le multiplicateur est de 2^e espèce ($\beta_\sigma > 0$) on a :

$$\alpha \beta = \omega^{\alpha_0 + \beta_0} \lambda_0 + \dots + \omega^{\alpha_0 + \beta_\sigma} \lambda_\sigma = \omega^{\alpha_0 \beta} ;$$

si le multiplicateur est de 1^{re} espèce, en vertu de la distri-

butivité tous les termes du multiplicande paraissent :
 $(\beta_\sigma = 0)$

$$\alpha \beta = \omega^{\alpha_0 + \beta_0} \lambda_0 + \dots + \omega^{\alpha_0 + \beta_{\sigma-1}} \lambda_{\sigma-1} + \omega^{\alpha_0} k_0 \lambda_\sigma + \omega^{\alpha_1} k_1 + \dots + \omega^{\alpha_\tau} k_\tau$$

Enfin pour l'exponentiation si l'exposant est de 2^e espèce :

$$\alpha^\beta = \omega^{\alpha_0 \beta}$$

si l'exposant est de 1^{re} espèce: $\beta = \beta' + \lambda_\sigma$, β' étant de 2^e espèce,

$$\alpha^\beta = \omega^{\alpha_0 \beta'} \cdot \alpha \lambda_\sigma$$

De la 2^e règle de la multiplication se déduit la décomposition univoque d'un nombre en ses facteurs premiers ; on peut, en effet, écrire α sous forme du produit :

$$\alpha = (\omega^{\alpha_1} k_1 + \dots + \omega^{\alpha_\tau} k_\tau) (\omega^{\alpha_0 - \alpha_1} k_0 + 1)$$

En itérant jusqu'à α_τ on obtient donc :

$$\alpha = \omega^{\alpha_\tau} k_\tau (\omega^{\alpha_{\tau-1} - \alpha_\tau} k_{\tau-1} + 1) \dots (\omega^{\alpha_0 - \alpha_1} k_0 + 1)$$

Et comme, suivant la même règle :

$$(\omega^\lambda k + 1) = (\omega^\lambda + 1) k$$

si on pose $\gamma_0 = \alpha_\tau$, $\gamma_1 = \alpha_{\tau-1} - \alpha_\tau$, etc., on obtient la décomposition canonique :

$$\alpha = \omega^{\gamma_0} k_\tau (\omega^{\gamma_1} + 1) k_{\tau-1} (\omega^{\gamma_2} + 1) k_{\tau-2} \dots (\omega^{\gamma_\tau} + 1) k_0$$

Si on ne retient des conditions pour l'exponentiation que 2) et 4) on définit ce que HAUSDORFF (1) devait appeler les *fonctions normales*. Elles se caractérisent par l'intéressante propriété de posséder des *nombre critiques*. Pour tout argument ξ d'une telle fonction on a :

$$f(\xi) \geq \xi$$

(sans quoi f déterminerait une application semblable de la

(1) HAUSDORFF (I), p. 114.

classe II sur elle-même où un élément serait appliqué sur un élément qui le précède). Les nombres critiques sont ceux pour lesquels :

$$f(\xi) = \xi$$

CANTOR n'étudie que la seule fonction γ_ξ — mais ses raisonnements valent pour toute fonction normale — : les nombres critiques correspondants sont les *nombres- ε* . Leur existence est démontrée par une construction (et résulte immédiatement des propriétés des suites fondamentales) : soit un nombre α qui ne soit pas critique :

$$f(\alpha) > \alpha$$

en posant $\alpha_1 = f(\alpha)$ $\alpha_2 = f(\alpha_1)$, etc., ... on obtient une suite fondamentale (propriété 2). Sa limite (propriété 4) est donnée par :

$$f(\beta) = f(\lim_{\nu} \alpha_{\nu}) = \lim_{\nu} f(\alpha_{\nu}) = \lim_{\nu} \alpha_{\nu+1} = \beta$$

β est le 1^{er} nombre critique supérieur à α . On voit immédiatement ⁽¹⁾ en effet qu'il ne peut y avoir d'autre nombre critique entre α et β , d'autre part que le 1^{er} nombre critique après β est obtenu par le même procédé en posant $\alpha = \beta + 1$, enfin que toute suite fondamentale de nombres critiques a pour limite un nombre critique. L'ensemble des nombres critiques relatifs à une fonction donnée s'obtient donc par deux procédés semblables aux deux principes d'engendrement des nombres de la classe II : il est bien ordonné et semblable à celle-ci. Pour la fonction γ_ξ , quel que soit γ , en vertu

(1) En effet, on aurait, comme $\alpha < \gamma < \beta$, $f(\alpha) = \alpha_1 < f(\gamma) = \gamma$ de même $\alpha_2 < \gamma$, etc., ... donc $\lim_{\nu} \alpha_{\nu} = \beta < \gamma$. Par suite, comme $\beta + 1$ suit immédiatement β , s'il y avait un nombre critique entre β et ce nouveau nombre, il serait également entre $\beta + 1$ et lui, ce qu'on vient de voir être impossible. Enfin soit la suite $\beta_1, \beta_2, \dots, \beta_{\nu}, \dots$ de nombres critiques, on sait qu'on a :

$$f(\lim_{\nu} \beta_{\nu}) = \lim_{\nu} f(\beta_{\nu}) = \lim_{\nu} \beta_{\nu}$$

donc $\lim_{\nu} \beta_{\nu}$ est un nombre critique.

de la forme canonique des nombres de la classe II, il n'y a qu'un ensemble de nombres critiques, dont on obtient le plus petit ε_0 en posant $\gamma = \omega : \varepsilon_0$ est évidemment la limite de la suite :

$$\omega, \omega^\omega, \omega^{\omega^\omega} \dots$$

Pour tout nombre α inférieur à un ε on a les trois égalités : (1)

$$\alpha + \varepsilon = \varepsilon \qquad \alpha\varepsilon = \varepsilon \qquad \alpha^\varepsilon = \varepsilon.$$

Telle est cette théorie des ordinaux transfinis — dernier étage de l'édifice cantorien — dont l'harmonie, la pureté de lignes devaient exciter l'admiration d'HILBERT. On a vu qu'elle découle tout entière — une fois admises les définitions intuitives de la somme et du produit de deux ensembles bien ordonnés — d'un approfondissement des relations établies par les deux principes générateurs des ordinaux. La simplicité des résultats vient de l'unité interne des méthodes. Mais de là aussi un divorce avec la théorie cardinale des ensembles, avec les problèmes posés initialement : l'application semblable est décidée par une loi ; la notion de puissance est purement descriptive ; il s'agit de prouver une corrélation biunivoque possible entre système de tous les objets jouissant des propriétés A et B . Sans doute la définition des ordinaux de la classe II permet-elle d'atteindre une nouvelle puissance : grâce au 3^e principe d'engendrement on a l'espoir de construire une échelle exhaustive des puissances. Mais ce principe, on l'a vu, n'est que négatif ; en réalité la définition de la suite des puissances obéit à deux nouveaux principes : le 1^{er} fondé sur le théorème que toute classe de tous les ordinaux dénotant les ensembles

(1) Soit α_0 le *degré* de α , le degré de $\varepsilon = \omega^\varepsilon$ est ε , donc α a un degré inférieur à ε , donc $\alpha + \varepsilon = \varepsilon$, de même $\alpha_0 + \varepsilon = \varepsilon$. D'autre part :

$$\alpha\varepsilon = \alpha\omega^\varepsilon = \omega^{\alpha_0 + \varepsilon} = \omega^\varepsilon = \varepsilon$$

donc aussi : $\alpha_0\varepsilon = \varepsilon$. Enfin :

$$\alpha^\varepsilon = \alpha^{\omega^\varepsilon} = \omega^{\alpha_0\omega^\varepsilon} = \omega^{\alpha_0\varepsilon} = \omega^\varepsilon = \varepsilon.$$

d'une certaine puissance α possède la puissance immédiatement supérieure ; le 2^e, transposition du 2^e principe d'engendrement des ordinaux : toute suite croissante infinie de nombres cardinaux détermine de façon univoque un nombre cardinal immédiatement supérieur à tout nombre de la suite (ainsi se définit $\aleph_\omega$). Mais l'utilité de l'échelle ainsi constituée n'apparaît que si l'on est sûr qu'y figure tout nombre cardinal, si d'autre part on a un moyen d'y fixer sa place. Or la théorie ordinale seule ne permet de répondre ni à l'une ni à l'autre question : l'universelle comparabilité, en tant que démontrée entre les ordinaux, ne servirait que si la 1^{re} question était résolue, loin qu'elle aide à la résoudre ; quant au 1^{er} principe, au lieu de permettre une assimilation entre $\aleph_1$ et le continu, il repose sur un procédé qui ouvre l'ère des paradoxes menaçant de ruiner toute la théorie.

En 1895, d'après le témoignage de BERNSTEIN, (1) CANTOR eut la première idée du paradoxe que BURALI-FORTI devait exposer en 1897 : l'ensemble de tous les nombres ordinaux (tels que les peuvent procurer les deux principes d'engendrement et le principe d'arrêt) rangés par ordre de grandeur forme un ensemble bien ordonné, qui a donc pour type un nombre ordinal α ; mais celui-ci ne peut figurer dans la suite Ω car à son rang il y dénoterait un segment de type α (l'ensemble de tous les ordinaux le précédant) et Ω serait semblable à un de ses segments. Antinomie qui oblige CANTOR à restreindre son mémoire de 1897 à la considération de la classe II, à laisser l'édifice inachevé. En 1899 quelques lettres adressées à DEDEKIND (2) essaient de la rendre inoffensive, bien plus de l'utiliser pour résoudre le problème de la comparabilité : « déjà depuis deux ans je suis

(1) BERNSTEIN (I), p. 187.

(2) Publiées in CANTOR (XXVIII), p. 443-450.

en possession d'une démonstration qu'il n'y a pas d'autre [puissance que les alephs] ; de sorte que par exemple un aleph déterminé correspond comme nombre cardinal au continu linéaire arithmétique.

Le principe est la distinction entre *multiplicités inconsistantes* « telles que l'hypothèse d'une réunion (ZUSAMMEN-SEIN) de tous leurs éléments conduit à une contradiction, de sorte qu'il est impossible de considérer la multiplicité comme une unité, une « chose finie », et *multiplicités consistantes*, non contradictoires en tant que *choses unes*. Les secondes sont les *ensembles* dont s'occupe la théorie. Exemples des premières sont le système Ω de tous les ordinaux transfinis et celui de tous les alephs, $\aleph$, parce que semblable à Ω . Deux systèmes équivalents sont ou tous deux consistants ou tous deux inconsistants ; toute partie d'un ensemble est un ensemble, tout ensemble d'ensembles est un ensemble. Enfin, si un système V n'a comme puissance aucun aleph « on reconnaît facilement qu'on peut projeter sur lui tout le système Ω , c'est-à-dire qu'il doit exister un système partiel V' de V qui est équivalent à Ω , V' est inconsistent » (1), donc V aussi ; donc tout système consistant a pour puissance un aleph. ZERMELO aperçoit dans le « on reconnaît facilement » un renvoi à l'intuition d'une bonne ordination transfinie : « visiblement CANTOR se représente les éléments de V successivement et arbitrairement coordonnés aux nombres de Ω , de sorte que tout élément de V n'est employé qu'une fois. Le procédé devrait ou bien se terminer une fois par l'épuisement de V , et V serait coordonné à un segment de Ω et sa puissance serait un aleph, contre l'hypothèse. Ou, au contraire, V reste inépuisable, donc contient une partie équivalente à Ω » (2). On verra pourtant que cette solution au problème de la

(1) *Ibid.*, p. 447.

(2) CANTOR (XXVIII), p. 451.

comparabilité et l'évasion corrélatrice de l'antinomie annoncent un travail analogue de VON NEUMANN ⁽¹⁾. Mais il fallait pour l'entreprendre abandonner les considérations intuitives. D'où l'échec de la tentative de 1897 ; un mois après sa première lettre, CANTOR fait lui-même l'objection décisive : « on doit soulever la question d'où je puis bien savoir que les multiplicités bien ordonnées... auxquelles j'attribue les nombres cardinaux :

$$\aleph_0, \aleph_1 \dots \aleph_{\omega_0}, \dots \aleph_{\omega_1} \dots$$

sont aussi réellement des... multiplicités consistantes. Ne serait-il pas concevable que déjà *ces* multiplicités soient inconsistantes et que la contradiction de l'hypothèse d'une « réunion » de tous leurs éléments ne se soit *seulement pas encore fait remarquer* ? Ma réponse est que cette question est également à étendre aux multiplicités finies... En d'autres termes : le fait de la « consistance » des multiplicités finies est une vérité simple indémontrable, c'est l'« axiome de l'arithmétique » (au sens ancien du mot). Et également la « consistance » des multiplicités auxquelles j'attribue les alephs comme nombres cardinaux est « l'axiome de l'arithmétique transfinie élargie » ⁽²⁾. »

A ces axiomes « au sens ancien » qu'aucune évidence ne garantissait plus, les successeurs de CANTOR s'efforceront, pour maintenir l'essentiel de la théorie abstraite, de substituer dans la mesure du possible la cohésion d'une axiomatique moderne. Comme point de départ pour leurs efforts se présentaient naturellement la construction systématique et les méthodes par lesquelles l'esprit algébriste de DEDEKIND avait ébauché un fragment de théorie des ensembles qui dût le moins possible à l'intuition.

(1) V. ci-après, p. 132.

(2) L. du 28 août 1899, *ibid.*, p. 447-448.

CHAPITRE III

Dedekind et la chaîne. Les axiomatisations.

« Il arrive fréquemment que dans un système donné Ω d'objets ou d'éléments, tout élément ω est remplacé, d'après une certaine loi, par un élément déterminé lui correspondant ; on a l'habitude d'appeler un tel acte une substitution... On s'exprime de façon encore plus commode si, comme nous le ferons, on considère cette substitution comme une application du système Ω et, de la sorte, appelle ω' l'image de ω , Ω' l'image de Ω . » « Sur cette capacité de l'esprit de comparer une chose ω avec une chose ω' , ou de mettre en relation ω avec ω' ... sans laquelle aucune pensée en général n'est possible, repose aussi, comme j'essaierai de le montrer ailleurs, toute la science des nombres ⁽¹⁾. » Tel est le premier énoncé publié de l'idée centrale de *Was sind und was sollen die Zahlen* ? probablement aussi — quoique les trois premiers brouillons soient antérieurs (échelonnés entre 1872 et 1878) — l'origine de l'entreprise, sous l'influence des travaux arithmétiques de DIRICHLET. La différence avec le point de vue de CANTOR est assez nette : il ne s'agit pas pour résoudre des problèmes mathématiques de développer une nouvelle science, mais de fonder la mathématique déjà existante. C'est une œuvre d'analyse des notions, de critique logique, non de création : DEDEKIND s'insère de lui-même ⁽²⁾ dans la suite, inaugurée par SCHRÖ-

(1) DIRICHLET (V). Supplément XI (rédigé par DEDEKIND), p. 470 et n. (paru en 1879).

(2) « La parution de ces mémoires [SCHRÖDER, *Lehrbuch der Arithmetik*

DER, KRONECKER et HELMHOLTZ — que devaient continuer FREGE et PEANO —, des essais de réduction des mathématiques à la logique. « Les nombres sont de libres créations de l'esprit humain, ils servent comme moyen pour saisir la diversité des choses plus facilement et avec plus de précision. » Telle est la réponse à la question du titre. Fonder « cette partie de la logique qui traite des nombres... [en n'admettant] rien de démontrable sans démonstration », tel est l'objet de l'ouvrage. Toutefois il ne s'agit pas d'« un essai de systématisation scolaire » comme ZERMELO devait caractériser plus tard le *Formulaire* de PEANO ⁽¹⁾, mais la griffe du mathématicien se montre dès la méthode choisie. Par « une décomposition raisonnée des suites de pensées sur lesquelles reposent les lois des nombres » ⁽²⁾, par la suppression de ces recours à l'« intuition intérieure » qui n'est qu'un raccourci, et surtout la systématisation de toutes les démarches élémentaires autour d'un acte unique, « relier des choses à d'autres », sera possible un travail de portée mathématique parce qu'on ne sort pas de l'activité mathématique. « Ἀεὶ ὁ ἄνθρωπος ἀριθμητίζει ... dès notre naissance... par cet exercice incessant déjà dans nos premières années..., nous acquérons un trésor de vérités proprement arithmétiques auxquelles plus tard nos premiers maîtres se réfèrent comme à quelque chose de simple, évident, donné dans l'intuition intérieure ⁽³⁾. » Ainsi cette expérience arithmétique, analysée conceptuellement, mais aussi « ramenée à une base unifiée » peut-elle servir de départ pour la création de notions nouvelles qui dépasseront l'arithmétique. Si « en opposition à CAN-

und der Algebra, KRONECKER *Über den Zahlbegriff*, HELMHOLTZ *Zählen und Messen* Leipzig, 1887] est l'occasion qui m'a déterminé à mettre en avant ma conception sous maints rapports semblable quoiqu'essentiellement différente pour le fondement. » DEDEKIND (II), t. III, p. 335.

(1) ZERMELO (II), p. 115.

(2) DEDEKIND (II), t. III, p. 337.

(3) *Ibid.*

TOR » ⁽¹⁾ on peut montrer que « maints concepts en réalité très complexes (comme par exemple celui du nombre [ANZAHL] des objets) sont tenus a tort pour simples », leurs arrangements ou leurs simplifications ne permettront-ils pas des extensions imprévues ?

Telle est la singularité décisive dans l'histoire de la théorie des ensembles, que ses notions, quelques uns de ses résultats essentiels aient été retrouvés, comme involontairement, au détour d'une recherche sur le fondement de l'arithmétique. Si le transfini, parce qu'échappant à l'évidence intuitive, réclame consolidation et définition du domaine mathématique, systématiquement la simple investigation critique de celui-ci pour le fini seul amène au transfini. Mais dans un cas c'est l'infini concret d'êtres saisissables et s'engendrant au sein d'une activité qui ne veut pas de bornes, dans l'autre c'est l'indétermination de l'abstraction. Avec DEDEKIND il n'y a plus d'ensembles effectifs de la réalité, mais systèmes de relations entre ensembles possibles quelconques et méthodes pour les définir. La seule analyse procure cependant trois résultats principaux : une organisation de la théorie générale autour de la notion d'application, une définition de l'ordre et corrélativement le fondement des définitions par récurrence, enfin la théorie complète des ensembles finis.

La définition du matériel de la théorie générale est simple précision du bien commun de cette époque. « Une *chose* est tout objet de notre pensée. » Lorsque « différentes choses *a*, *b*, *c*... conçues d'un même point de vue, sont réunies dans l'esprit, on dit qu'elles forment un *système S* ». Les choses *a*, *b*, *c*... sont les éléments du système qui « comme objet de notre pensée est également une chose ». Enoncé essentiel,

(1) Note marginale sur le premier brouillon de la préface (datée : 8 juillet 1887).

mais déjà fait par BOIZANO et CANTOR : un système S « est complètement déterminé si, pour toute chose, il est déterminé qu'elle soit ou non élément de S ». Donc deux systèmes qui ont mêmes éléments sont identiques ; on distingue évidemment entre un élément et le système formé de ce seul élément. Cependant, malgré les hésitations à ce sujet, dont les brouillons de 1872 à 1878 portent trace, le même signe dénote la relation ensemble-partie et ensemble-élément ⁽¹⁾. Ou plutôt la seconde semble reléguée en dehors de la théorie ; d'où, à la fin surtout, quelques flottements. C'est qu'au départ au moins, comme les éléments abstraits ne sont pas distinguables, seules les relations entre ensembles sont objets d'étude : réunion, intersection et, avant tout, application. Pour celle-ci la différence avec l'attitude cantorienne apparaît nettement : il ne s'agit pas d'un simple élargissement de l'opération géométrique de la projection — mise en correspondance de deux réalités données — mais d'une généralisation du rapport fonctionnel. Deux conséquences : 1° l'application n'est pas forcément biunivoque (ce qu'elle était chez CANTOR puisqu'elle ne servait qu'à définir la puissance) ; 2° surtout la seule application considérée — étudiable *in abstracto* — sera l'application d'un système en lui-même par rapport à quoi il est une *chaîne*. Quel autre mode de séparation, en effet, entre ensembles, indiscernables puisque abstraits, que la relation de partie à tout. A l'exploration des propriétés de la *chaîne* est consacré tout l'ouvrage.

En premier lieu, fondement sur elle du raisonnement d'induction complète. Le caractère de chaîne — relativement

(1) Nous emploierons les signes modernes : $m \in M$: m est élément de M ; $N \subset M$: N est *partie* de M ; $M + N$: *réunion* (système de tous les éléments contenus dans M et N) de M et N ; $M.N$ *intersection* de M et N (partie commune). En principe le système formé de l'élément a sera $\{a\}$. Pratiquement, quand il n'y aura pas d'ambiguïté, le système sera aussi désigné par a . Ainsi la *chaîne* du système $\{a\}$ sera désignée par a_0 au lieu de $\{a\}_0$.

à une application donnée φ — est invariant par rapport à cette application (l'image d'une chaîne est une chaîne), la réunion et l'intersection. D'où la possibilité de définir une sorte de système maximum ou d'enveloppe d'un sous-ensemble relativement à l'application : c'est la *chaîne propre* ⁽¹⁾ M_o , intersection de toutes les chaînes dont M est partie. Il importe de noter le caractère transfini — c'est-à-dire abstrait — de la définition : la totalité des chaînes contenant M n'est donnée par aucun caractère, mais purement posée *in abstracto*. Le rôle maximal s'en déduit, qui apparaît mieux avec la propriété caractéristique de la *chaîne-image* : l'image de la chaîne propre d'un ensemble est la chaîne propre de son image ⁽²⁾ :

$$(M')_o = (M_o)' = M'_o$$

On en tire la décomposition ⁽³⁾ :

$$M_o = M + M'_o$$

M'_o représente donc le maximum de ce que l'application itérée φ permet d'atteindre à partir de M (il y a chaque fois débordement sur le domaine base puisque M n'est pas une chaîne). Le principe d'induction complète se formule alors : si une propriété est vraie pour M et si l'application φ la transmet, elle est aussi vraie pour la chaîne propre M_o de M par rapport à φ . La démonstration est immédiate à partir des définitions et propriétés précédentes ⁽⁴⁾.

(1) Pour abrégier — sauf dans les cas prêtant à la confusion — nous disons chaîne de M .

(2) En effet $M \subset M_o$ donc $M' \subset (M_o)'$ or $(M_o)'$ est une chaîne donc $(M')_o \subset (M_o)'$; inversement $M' \subset (M')_o$ donc $M + (M')_o$ est une chaîne, donc $M_o \subset (M')_o + M$ donc :

$$(M_o)' \subset [(M')_o]' + M' \subset (M')_o.$$

(3) On a $M \subset M_o$ et puisque $M' \subset M_o$: $(M')_o \subset M_o$; avec l'inégalité de la note précédente on en tire l'égalité.

(4) En effet : qu'une propriété soit vraie de M s'écrit $M \subset \Sigma$ (Σ = système des objets possédant la propriété). Par hypothèse, si $G = M_o$, Σ , on a $G' \subset \Sigma$, donc $G' \subset G$, donc G est une chaîne, donc $M_o \subset G$ donc $M_o' \subset G \subset \Sigma$.

Telle est la théorie minima, celle qui ne dépend que des notions corrélatives d'ensemble et d'application (en général, c'est-à-dire non biunivoque). Pour aller plus loin il est immédiat de faire intervenir l'application biunivoque et formuler une hypothèse sur le comportement à son égard de la relation partie-tout : ou bien un ensemble peut être appliqué de façon biunivoque sur une vraie partie de lui-même, ou la chose est impossible. Bifurcation entre ensembles finis et infinis qui s'imposait avec la notion d'application et à laquelle DEDEKIND prétend avoir été amené par ses propres réflexions : « plusieurs années avant la parution du mémoire de CANTOR et à une époque où le nom même de l'œuvre de BOLZANO m'était complètement inconnu » ⁽¹⁾. Pour inaugurer la théorie des ensembles infinis il faut un théorème d'existence ; ici la rencontre avec BOLZANO se précise : l'exemple d'ensemble infini sera « le monde de mes pensées » dont chacune possède une image, la pensée de cette pensée, application évidemment biunivoque. BOLZANO avait, pour prouver qu'il y a, « dans le règne des choses qui sont simplement possibles..., incontestablement des ensembles infinis », invoqué « l'ensemble des propositions et vérités en soi » ⁽²⁾. Sa définition de l'infini est — on l'a vu — pourtant différente : pour toute vérité, nous en pouvons énoncer une autre, savoir que la première est vraie ; ainsi avons-nous une loi permettant « de déduire de la proposition *A* celle-ci différente d'elle... *B*, et de *B* à nouveau *C* et ainsi sans fin ». Cette suite de propositions est donc semblable à la suite des nombres, c'est-à-dire « possède une multiplicité plus grande que tout nombre, donc est infinie ». Mais le principe est le même, expliquant non seulement la rencontre, mais le recours à l'univers psychologique, étrange dans un écrit comme celui de DEDEKIND :

(1) *Ibid.*, p. 342.

(2) BOLZANO (II), p. 14.

il s'agit du pouvoir indéfini de thématization décrit de façon analogue par les deux auteurs (faire d'une proposition — ou pensée — le sujet d'une autre proposition, ou pensée) et source véritable de l'infini mathématique. L'itération simple n'en est qu'un cas particulier, dans la numération, par exemple, où la collection finie des objets déjà numérotés est rassemblée dans un système, analogue à l'unité, à laquelle donc on peut adjoindre l'unité. Comme on voit, l'acte fécond n'est pas l'adjonction, mais l'acte de rassemblement préalable. Lorsque, en 1911, après huit ans d'hésitations, DEDEKIND se décide à rééditer son livre malgré « l'importance et la justesse partielle des doutes... exprimés sur la solidité de ses fondements », doutes qui portaient en premier lieu contre la démonstration de l'infini, il se borne en justification à affirmer sa confiance « qu'une investigation rigoureuse de la puissance créatrice de l'esprit capable de faire à partir d'éléments déterminés, un nouvel élément, leur système, qui est nécessairement différent de chacun de ces éléments, doit conduire à reconstituer de façon inattaquable les fondements de mon travail » (1). Former un système, par exemple, à partir d'un élément le système de cet élément, c'est bien thématizer, rendre une pensée sujet (logique) d'une autre pensée. Le théorème de l'infini est partie essentielle de l'œuvre, justification unique aussi d'un infini mathématique dans n'importe quelle théorie. Mais BOLZANO ne prouvait qu'un infini virtuel: ce n'est que par la notion d'un en soi des propositions qu'il pouvait confondre dans sa définition infini syncatégorématique — ou indéfini — avec infini catégorématique. Il est remarquable que la démonstration de DEDEKIND ne prouve pas davantage: pour atteindre l'infini par l'application il faudrait que le système des pensées de pensées soit vraie partie du système des pensées; or c'est « mon propre moi » (2) qui est

(1) Préface de la 3^e éd. DEDEKIND (V), p. 343.

(2) *Ibid.*, p. 357. Cf. L. inédite à KEFERSTEIN (8-2-1890). « Je tiens la

exemple d'une pensée qui ne serait pas pensée de pensée. Seul moment extra-mathématique du raisonnement, mais décisif, chute dans ce que HILBERT appelait le transcendantal, ou réalisme psychologique. En fait, la reconnaissance du pouvoir de thématization ne permet d'affirmer qu'un infini défini par l'ordre. La propriété prise comme départ par DEDEKIND étant plus forte, permet de déduire l'ordre.

On appelle système simplement infini un système N qui soit la chaîne propre d'un de ses éléments par rapport à une application biunivoque, ϕ , de N sur une de ses vraies parties. L'élément ainsi distingué — élément fondamental — n'appartient évidemment pas à l'image. Tout système infini contient un système simplement infini : il suffit de prendre la chaîne d'un des éléments extérieurs à l'image dans une des applications transformant le système en une de ses vraies parties. Un système simplement infini est ordonné par l'application qui le définit comme tel ; en effet, d'une part tout élément est distinct de son image, d'autre part étant donnés deux éléments quelconques m et n , on a entre leurs chaînes propres et leurs chaînes images une et une seule des 2 relations :

$$\begin{array}{ccc} n_o \subset m'_o & m_o \subset n'_o & \\ \text{qui s'écriront} & & \\ m < m & n < m & \end{array}$$

On voit que les caractéristiques de l'ordre sont présentes : asymétrie (les 2 relations précédentes s'excluent) ; non réflexivité : il est impossible que l'on ait :

$$n_o \subset n'_o ;$$

enfin transitivité :

$$n_o \subset m' \quad \text{et} \quad m_o \subset p' \quad \text{entraînent} \quad n_o \subset p'.$$

démonstration pour rigoureusement correcte : mon moi peut être objet de ma pensée... mais n'est pourtant pas lui-même une pensée. »

On voit immédiatement que tout élément précède, dans l'ordre, son image. Cet ordre est un bon ordre : en effet toute chaîne du système est la chaîne propre d'un élément ; considérons la chaîne T_0 d'une partie quelconque T du système : T_0 étant la chaîne de l'élément k on a :

$$T_0 = k + k'_0$$

T contient k (sans quoi la chaîne de T serait k'_0) et tout élément de T distinct de k est contenu dans K'_0 , donc le suit dans l'ordre ; toute partie T a un premier élément. L'élément fondamental du système tout entier précède tous les autres : on le désigne par 1 ; les éléments du système sont les nombres ordinaux naturels ; le système entier, la suite N de ces nombres. La notion de segment Z_n , ensemble de éléments $\leq n$, n étant un élément donné, permet de retrouver les propriétés classiques (tout nombre a , dans l'ordre, un prédécesseur et un suivant immédiat, toute partie d'un segment possède un plus grand nombre ; toute partie du système qui ne possède pas de plus grand nombre est simplement infinie). Le raisonnement employé en général pour les démonstrations est l'induction complète qui revêt ici son aspect habituel. Pour qu'une propriété soit vraie de tous les éléments de la chaîne $1_0 = N$, il suffit de la démontrer pour 1, et de prouver que l'ordre (l'application φ) la transmet d'un élément à son successeur immédiat (son image).

La conséquence immédiate et la plus importante de cette théorie de l'ordre est la justification des définitions par récurrence, apport original de DEDEKIND. Etant donné un système quelconque Ω appliqué sur lui-même par l'application θ (biunivoque ou non), une définition par récurrence consiste à établir une correspondance entre N et Ω telle que :

1° l'image de 1 soit un élément donné ω de Ω ;

2° l'image de tout nombre n' soit l'image par rapport à θ de l'image dans Ω de son prédécesseur n . Il y a une sorte de transport de la mise en ordre φ sur l'application θ . De telles définitions sont justifiées si l'on démontre qu'étant donnés ω et θ il existe une et une seule application Ψ satisfaisant aux 3 conditions :

1. $\Psi(N) \subset \Omega$
2. $\Psi(1) = \omega$
3. $\Psi(n') = \theta \bar{\Psi}(n)$

La démonstration s'effectue grâce au *procédé de fusionnement*. On prouve d'abord que pour tout nombre n il existe une et une seule fonction Ψ_n satisfaisant aux 3 conditions :

- 1'. $\Psi_n(Z_n) \subset \Omega$
- 2'. $\Psi_n(1) = \omega$
- 3'. $\Psi_n(m') = \theta \Psi_n(m)$ pour tout $m < n$.

Par induction complète : la fonction Ψ_1 existe évidemment (il suffit de la définir par la propriété 2' : 1' est alors satisfaite, 3' disparaît). Si la fonction Ψ_n existe et est déterminée de façon univoque, on construit $\Psi_{n'}$, en posant :

$$\begin{aligned} \Psi_{n'}(m) &= \Psi_n(m) \quad \text{pour tout } m \leq n \\ \text{et} \quad \Psi_{n'}(n') &= \theta \Psi_n(n) \end{aligned}$$

$\Psi_{n'}$ est complètement définie et satisfait aux conditions 1' — 3'.

Enfin on obtient Ψ en écrivant :

$$\Psi(n) = \Psi_n(n)$$

Définition évidemment univoque et garantissant l'existence (puisque pour tout n il existe un et un seul Ψ_n).

Ainsi s'introduisent les opérations élémentaires si on prend pour système Ω le système N lui-même ; pour l'addition de $+$ $m : n + m$ (n étant un nombre choisi) on a :

$$\omega = m \quad \theta(n) = n'$$

Pour la multiplication par m , on pose :

$$\omega = m \quad \theta(n) = n + m$$

Enfin, l'exponentiation du nombre a :

$$\omega = a \quad \theta(n) = n \times a$$

Les propriétés classiques sont immédiatement retrouvées par induction complète. Il importe de noter, comme fait DEDEKIND, le caractère spécial des définitions par récurrence : elles ne sont fondées que par rapport au système simplement infini N (si on suppose sa non contradiction comme système ordonné $1_0 = N$) ; tandis que l'induction complète est utilisable, comme on l'a vu, pour toute chaîne relative à une application quelconque (biunivoque ou non). Si au lieu de N on prend une telle chaîne, le procédé de définition par récurrence aboutit, en général, à des contradictions ⁽¹⁾.

La théorie précédente a son utilisation immédiate dans la théorie des groupes : c'est grâce à elle que prend un sens la notion de $n^{\text{ième}}$ itération d'une opération. En prenant un système Ω d'objets quelconques où la liaison de deux éléments ω et ν est déterminée (représentée par le symbole $\omega\nu$), on peut considérer que tout élément choisi au hasard, ω , détermine une application de Ω en lui-même (à tout ν correspond $\omega\nu$). L'application n fois itérée de Ω par ω sera définie par les conditions :

$$\text{II. } \omega' = \omega$$

$$\text{III. } \omega^{n'} = \omega \omega^n$$

(1) DEDEKIND prend l'exemple du système S composé de 2 éléments a et b , appliqué sur lui-même : $a' = b$ $b' = a$; donc $a_0 = b_0 = S$. Si on choisit pour Ω le système formé des 3 éléments α, β, γ avec l'application θ telle que :

$$\theta(\alpha) = \beta, \theta(\beta) = \gamma, \theta(\gamma) = \alpha$$

on voit qu'il est impossible de satisfaire aux conditions 1-3 : si $\Psi(a) = \alpha$, on aura :

$$\Psi(b) = \beta \text{ et } \Psi(b' = a) = \gamma \neq \alpha.$$

elle résulte donc d'une application (suivant le procédé de récurrence) de la suite des nombres N sur Ω . Ainsi peut s'interpréter la notion de chaîne propre : si l'on considère le système Ω de toutes les applications d'un système S en lui-même, les conditions II et III définissent bien toutes les itérations d'une certaine application ω ; elles ont, en effet, la forme des conditions 2 et 3 d'une définition par récurrence si on pose $\theta = \omega$; or ceci est possible en vertu des propriétés élémentaires de la chaîne : si

$$\omega(S) \subset S \quad (\text{vrai par hypothèse})$$

on voit que :

$$\omega[\omega(S)] \subset \omega(S)$$

Donc la liaison $\omega \omega$ fait passer d'un élément de Ω à un autre élément de Ω (application de S en lui-même), $\omega(v)$ (v étant un élément de Ω) est donc une application de Ω en lui-même, définissant θ . Considérons maintenant la chaîne propre $\omega_0(A)$ d'une partie quelconque A de S ; on voit que quel que soit n :

$$\omega^n(A) \subset \omega_0(A)$$

Si l'on pose :

$$B = \sum_n \omega^n(A) + A$$

d'une part, donc :

$$B \subset \omega_0(A)$$

d'autre part, B est une chaîne par rapport à ω ; en effet :

$$\omega(B) = \omega(A) + \sum_n \omega^n(A) \subset \sum_n \omega^n(A)$$

donc, en vertu de la définition de la chaîne propre :

$$\omega_0(A) \subset B \quad \omega_0(A) = B$$

Ainsi l'utilisation de la chaîne propre coïncide avec le procédé cantorien du passage à la limite ⁽¹⁾ : prendre l'intersec-

(1) D'où la coïncidence pour le fond entre la démonstration BERNSTEIN SCHRÖDER du théorème de l'équivalence et celle que Zermelo, et, sans qu'il le sache, avant lui DEDEKIND (V. CANTOR-DEDEKIND (I), p. 8 ; CANTOR-

tion de toutes les chaînes revient à considérer ce qui résulte de toutes les itérations d'une même application. Le même recours au transfini est impliqué. Il importe de remarquer que si la présentation de DEDEKIND est plus simple — puisque partant de principes préalables à la définition de la suite des entiers (et procurant celle-ci), — elle ne permet le passage aux autres présentations que si on considère comme démontré le théorème d'existence d'un ensemble infini (*sur quoi repose la non contradiction de la définition d'un ensemble simplement infini*). La notion de chaîne propre serait donc plus générale (puisque indépendante de ce théorème d'existence). Il est vrai que la non contradiction de sa définition devait elle-même apparaître plus tard problématique.

(XXVIII), p. 449, avec les remarques de Zermelo p. 450, enfin DEDEKIND (V), t. III, p. 447) obtiennent par la notion de chaîne. La démonstration DEDEKIND s'énonce ainsi, avec les notations déjà employées : si $M \infty P$ avec :

$$P \subset T = P + Q \subset M = P + Q + R$$

on doit avoir $M \infty T$.

Par rapport à l'application ϕ de M sur P , soit R_0 la chaîne de R , posons

$$M = R_0 + V;$$

on a, puisque $P = M'$:

$$P = R'_0 + V'$$

il en résulte :

$$V \subset T$$

$$(\text{puisque } R \subset R_0 \text{ et } R_0 + V = T + R)$$

et $R'_0 + V \subset T$

$$(\text{car } R'_0 \subset T, \text{ puisque } P = R'_0 + V' \subset T);$$

comme d'autre part :

$$T \subset R_0 + V = M = R + R'_0 + V,$$

et que T est par hypothèse extérieur à R , on a :

$$T \subset R'_0 + V \quad \text{donc} \quad T = R'_0 + V$$

Donc, puisque :

$$R'_0 \infty R_0 \text{ (ou } V' \infty V) : T \infty M \infty P.$$

Si l'on pose maintenant :

$$R_0 = R + R' + R'' + \dots \quad (R'' = \phi \phi R, \text{ etc...})$$

on a :

$$M = V + R + R' + R'' + \dots$$

$$P = V' + R' + R'' + \dots \subset T$$

donc, puisque R est extérieur à T , tous les $R', R'', \dots$ sont extérieurs à R ; on peut écrire $M - R = T = V + R' + R'' + \dots$ ce qui est évidemment équivalent à P et à M .

Le même procédé du *fusionnement* permet de démontrer la proposition capitale de la théorie des ensembles finis : la condition nécessaire et suffisante pour qu'un ensemble S ne soit pas fini est que tout segment Z_n de la suite des nombres soit applicable sur lui de façon biunivoque. La condition est évidemment nécessaire, puisque tout ensemble infini contient un ensemble simplement infini (donc applicable de façon biunivoque sur la suite N). La démonstration qu'elle est suffisante se fait par la construction d'une suite d'applications Ψ_n biunivoques définies respectivement pour chaque Z_n et telles que chacune d'elles soit contenue dans (coïncide avec) toutes les applications de numéro d'ordre plus élevé. Autrement dit pour $m < n$ et pour tout $t \leq m$ ($t \in Z_m$) on a :

$$\Psi_n(t) = \Psi_m(t)$$

La définition des Ψ_n est faite par récurrence ⁽¹⁾ grâce aux applications α_n données des Z_n . On en déduit une application biunivoque de toute la suite N en posant :

$$\chi(n) = \Psi_n(n)$$

on a bien :

$$\chi(N) \subset S$$

donc S est infini. L'exemple montre la symétrie entre le procédé de fusionnement et le procédé de la diagonale pour les systèmes ordonnés : dans les deux cas définition d'une nou-

(1) Soit Ω le système de toutes les applications biunivoques de tous les systèmes Z_n sur S : à toute application β d'un système Z_n on fait correspondre une application γ de $Z_{n'}$ en posant $\gamma(p) = \beta(p)$ pour tout $p \in Z_n$ et $\gamma(n') = \alpha_{n'}(k)$, $\alpha_{n'}$ étant l'application, donnée dans l'hypothèse, de $Z_{n'}$ sur S , k étant le plus petit nombre pour lequel l'image donnée par $\alpha_{n'}$ n'appartient pas à $\beta(Z_n)$ (il en existe évidemment puisque ces applications sont biunivoques, et que Z_n et $Z_{n'}$ ne peuvent être mis en correspondance biunivoque). Cette correspondance bien définie entre β et γ applique donc Ω sur lui-même (on voit que γ est biunivoque), c'est l'application θ des définitions par récurrence ; l'élément ω sera α_1 : chacune des valeurs de la fonction récurrente ainsi définie est un ψ_n .

velle fonction d'après l'ensemble des valeurs des fonctions d'une suite pour l'argument égal à leur numéro d'ordre, ici unissant toutes les fonctions, là distincte d'elles toutes. Du théorème découle que tout ensemble fini T est applicable de façon biunivoque sur un et un seul Z_n : le nombre n déterminant le segment sera le *nombre cardinal* de T .

Ainsi, d'une part la théorie cardinale se trouve dans la dépendance de la théorie ordinale, d'autre part celle-ci n'est établie que grâce à la considération de l'ensemble simplement infini : elle fonde le fini sur l'infini. Sans doute est-il possible, à partir de la théorie générale de l'application, de déduire déjà deux propriétés : toute partie d'un ensemble fini est finie, l'ensemble formé par adjonction d'un élément à un ensemble fini est fini ; en revanche, les propriétés caractéristiques comme : l'ensemble des parties d'un ensemble fini est fini, exigent le recours à l'ordre. Caractère boiteux de cette théorie (tolérable parce qu'elle n'est qu'un accessoire dans l'ouvrage dont le but est l'étude de la suite indéfinie des nombres) que DEDEKIND a lui-même senti : une autre définition du fini, basée également sur l'application et la chaîne, permet de déduire l'ordre sans passer par l'infini.

C'est l'ordre « cyclique » : un ensemble est fini lorsqu'il existe une application ϕ de l'ensemble sur lui-même telle qu'aucune des vraies parties de l'ensemble ne soit appliquée sur elle-même ⁽¹⁾. La notion de chaîne se transpose en *quasi chaîne* : toute partie H_b de T telle que l'image de tout élément de H_b soit contenue dans H_b sauf de celle d'un de ses éléments b ; à la chaîne propre correspond l'*arc* $a \ b$: in-

(1) Préface de la 2^e édition DEDEKIND (II), t. III, p. 342. Dans un fragment resté inédit de son vivant (DEDEKIND (V), t. III, p. 450-458), DEDEKIND avait commencé à développer les conséquences de cette définition, en particulier donné la notion d'arc et défini l'ordre : on peut en déduire les propriétés ci-dessus (J. CAVAILLÈS (I), p. 143-147) ; la démonstration que toute partie d'un ensemble fini est finie s'effectue en particulier sans recours à l'induction complète.

tersection de toutes les quasi chaînes H_b contenant l'élément a . On voit que $a b$ est également une quasi-chaîne et, grâce à cette notion, il est facile de démontrer que l'application φ est biunivoque, tout élément distinct de son image, enfin que — exactement comme la chaîne propre — un arc se décompose :

$$a b = a + a' b$$

Etant donnés deux éléments a et b , on a une et une seule des relations :

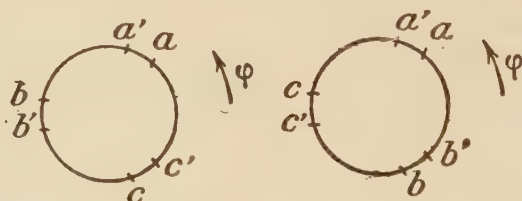
$$a \in a' b' \quad \text{ou} \quad b \in a' b'$$

qui permet de définir l'ordre ⁽¹⁾ :

$$a > b \quad \text{ou} \quad b > a$$

avec les propriétés habituelles. De là se déduisent le principe d'induction complète, les propriétés que toute partie est finie, que l'ensemble de toutes les parties est fini et qu'il est impossible d'appliquer de façon biunivoque un ensemble fini sur une de ses vraies parties. La réciproque de la dernière proposition, qui établirait l'équivalence des deux définitions, ne semble pas pouvoir être démontrée sans que l'on fasse intervenir, sous une forme ou une autre, le transfini.

Tel est le phénomène significatif auquel aboutit la théorie des ensembles finis : toutes les autres définitions adop-



(1) Nous simplifions; en réalité on définit du même coup l'ordre entre 3 éléments, a, b, c : on a :

$$\text{ou bien } b'c = b'a + a'c \text{ et } c'b = c'a.a'b$$

$$\text{ou bien } b'c = b'a.a'c \text{ et } c'b = c'a + a'b$$

tées ultérieurement, définition ZERMELO ⁽¹⁾ analogue à la 2^e définition de DEDEKIND, définition RUSSELL-WHITEHEAD ⁽²⁾ par le principe ordinaire d'induction complète (un ensemble A est fini s'il appartient comme élément à toute classe K contenant l'ensemble vide et telle que si $B \in K$ et $a \in A$, $B + a \in K$), définitions TARSKI ⁽³⁾ et JANISZEWSKI par les éléments *saturé* et *irréductible* (un ensemble A est fini si tout ensemble M non vide de ses parties contient un élément saturé, c'est-à-dire qui ne soit partie d'aucun autre élément de M , ou irréductible, c'est-à-dire dont aucune partie ne soit élément de M) sont équivalentes entre elles, ainsi que l'a montré TARSKI ⁽⁴⁾. Or le moyen employé uniformément pour démontrer les équivalences est l'induction complète : il semble donc que, sous-jacent à chacune de ces définitions, se trouve un renvoi à la définition arithmétique ordinaire, la construction progressive de l'ordre, intuition irremplaçable, comme le voulait Poincaré. Seule la propriété prise par DEDEKIND comme 1^{re} définition est pure de toute notion d'ordre : elle est démontrable à partir des autres définitions, mais ne permet pas de démontrer les propriétés énoncées dans celles-ci, à moins que l'on n'ait recours à un instrument sur lequel concentraient leurs attaques au début du siècle la plupart des mathématiciens hostiles à la théorie des ensembles, l'axiome de choix, principe qui se révélait fondamental pour l'étude abstraite du transfini.

On en trouve, note FRAENKEL ⁽⁵⁾, la première formule dans un écrit de PEANO de 1890 sur les équations différen-

(1) ZERMELO (IV).

(2) RUSSEL-WHITEHEAD (I).

(3) TARSKI (I) où se trouvent passées en revue et comparées les différentes définitions.

(4) *Ibid.*

(5) FRAENKEL (XI), p. 33.

tielles : « mais comme on ne peut pas appliquer une infinité de fois une loi arbitraire par laquelle à une classe a on fait correspondre un individu de cette classe, on a formé ici une loi déterminée par laquelle à chaque classe a , sous des hypothèses convenables, on fait correspondre un individu de cette classe ». En 1902, BEPPO LEVI l'énonce comme principe séparé à l'occasion d'une démonstration de BERNSTEIN ⁽¹⁾. En 1904, sur la suggestion d'Erhard SCHMID, ZERMELO le prend comme point de départ — et pierre angulaire — pour sa démonstration que tout ensemble peut être bien ordonné. Aussitôt apparaissent les objections, à la tête desquelles, par une curieuse rencontre, celle de PEANO. La critique de BOREL est particulièrement claire. Si la démonstration ZERMELO prouve l'équivalence entre les deux problèmes : mettre un ensemble M sous forme bien ordonnée, et faire correspondre à toute partie M' de M un élément distingue, on ne peut considérer qu'elle résolve le premier puisqu'elle laisse le second intact. « Un tel raisonnement ne me paraît pas mieux fondé que le suivant : pour bien ordonner un ensemble M il suffit d'y choisir arbitrairement un élément auquel on attribuera le rang 1, puis un autre auquel on attribuera le rang 2 et ainsi de suite *transitivement*... De tels raisonnements sont en dehors des mathématiques ⁽²⁾. » Ils reviennent, en fait, à supposer le bon ordre déjà donné dans l'ensemble considéré : il n'y a pas raisonnement mais cercle vicieux.

La réponse de ZERMELO est triple : 1^o le problème spécial posé n'était pas : donner un moyen de bien ordonner tout ensemble, mais démontrer que tout ensemble peut être mis sous forme bien ordonnée, problème spécifique du cantorisme qui, dans la théorie cardinale, part d'en-

(1) ZERMELO (II), p. 113 n.

(2) BOREL (II), p. 195.

sembles donnés concrètement comme le continu et veut effectuer la synthèse avec le développement conceptuel de la théorie ordinale.

2^o Une solution effective est donnée à ce problème grâce à une utilisation de la théorie abstraite de DEDEKIND qui remplace justement un procédé de *construction* transfini par une détermination *abstraite* de l'ordination. Il y a gain puisque dans le 1^{er} cas on ne peut assurer que l'ensemble *tout entier* est épuisé par la mise en ordre progressive, ce qui est évidemment le seul point en question. La démonstration elle-même se borne à utiliser le procédé de *fusionnement* pour les parties bien ordonnées de l'ensemble M : si l'on suppose coordonné à toute partie de M un *élément distingué*, il suffit d'exiger des parties bien ordonnées M_γ que chacun de leurs éléments soit élément distingué du complémentaire $M - A$ du segment A qu'il y détermine. Il existe de telles parties : l'ensemble $\{m\}$ formé de l'élément distingué de M . Etant données deux telles parties, l'une est identique à un segment de l'autre (avec même bon ordre) ; la réunion de toutes ces parties constitue un ensemble bien ordonné auquel appartient tout élément de M : conséquences immédiates de la notion de bon ordre. Le seul reproche que l'on pourrait faire au raisonnement est qu'il suppose établie déjà la théorie du bon ordre, c'est-à-dire non contradictoire la notion d'ensemble bien ordonné transfini : l'exemple prouvant qu'il existe des parties M_γ est un ensemble bien ordonné fini. Dans cette mesure l'objection de BOREL est justifiée : ce qu'il faut montrer — et qui n'apparaît pas ici — c'est que la théorie du bon ordre peut être fondée avec les seules ressources de la théorie abstraite des ensembles (et l'application).

D'où : 3^o la dernière réponse de ZERMELO ⁽¹⁾, une nou-

(1) ZERMELO (II), p. 107.

velle démonstration qui ne suppose préalablement fondées que les notions d'ensemble (avec les opérations réunion intersection), de partie d'ensemble et d'application, filiation cette fois directe à partir de DEDEKIND. « Un ensemble M est dit bien ordonné si à chacun de ses éléments a correspond de façon univoque une partie $R(a)$ de M , son *reste*, et si toute partie non vide P de M contient un et un seul élément p dont le reste $R(p)$ contient P », p est dit le *premier* élément de P . On voit qu'il y a ainsi un ordre entre les éléments de M puisque pour tout couple $\{m, n\}$ l'un des éléments est le *premier* des deux : on écrit $m < n$, relation équivalente ⁽¹⁾ avec $n \in R(m)$, qui exclut $m \in R(n)$; donc relation transitive : si $m < n < p$, on a forcément ⁽¹⁾

$$\{m, n, p\} \subset R(m) \text{ donc } p \in R(m) : m < p ;$$

enfin dans une partie quelconque le premier élément précède bien tous les autres : la définition habituelle est retrouvée. Le théorème de ZERMELO s'énonce alors : tout choix dans M (c'est-à-dire toute coordination à chaque partie A de M d'un élément *distingué* a de A) entraîne une bonne ordination de M , savoir celle où le reste $R(p)$ d'un élément p a pour élément distingué p . En effet, considérons une application φ de l'ensemble $\mathfrak{M}$ des parties de M en lui-même, faisant correspondre à toute partie P la partie $P' = P - \{p\}$ obtenue par retranchement de l'élément distingué de P , on appellera *θ-chaîne* par rapport à φ , toute chaîne telle que l'intersec-

(1) En effet : 1° $m < n$, d'après sa définition, entraîne $\{m, n\} \subset R(m)$ donc $n \in R(m)$; 2° si $n \in R(m)$ on a $\{m, n\} \subset R(m)$ et m est le *premier* des deux éléments.

2° Par définition $\{m, n, p\}$ est partie d'un des trois restes $R(m)$, $R(n)$, $R(p)$; or les deux autres inclusions : $\{m, n, p\} \subset R(n)$ et $\{m, n, p\} \subset R(p)$ sont impossibles par hypothèse (puisque $m < n$ on ne peut avoir $m \in R(n)$, etc...)

tion d'un système quelconque de ses éléments lui appartient. Soit M_0 la θ -chaîne propre de $\{M\}$, on voit que pour tout couple d'éléments N et P de M_0 , on a ou $N \subset P$ ou $P \subset N$ (en effet le système des éléments jouissant de cette propriété dans M_0 constitue une θ -chaîne propre contenant M). Etant donnée une partie quelconque Q de M , appelons *enveloppe* Q_1 de Q l'intersection de tous les éléments de M_0 qui contiennent Q (il en existe au moins un, M), Q_1 est le reste $R(q)$ de l'élément distingué de Q . En effet, l'élément distingué de Q , est nécessairement dans Q (sans quoi $Q_1 \subset Q$ élément de M_0 , contiendrait aussi Q) et Q_1 est le seul élément de M_0 à jouir de la propriété ; soit d'autre part Z l'enveloppe de $\{q\}$, on a évidemment $Z \subset Q_1$ (puisque Z est élément de M_0), considérons $Q' = Q - \{q\}$, l'enveloppe de Q' est l'image Q'_1 , de Q_1 dans l'application φ , comme Q'_1 et Z sont éléments de M_0 on a $Q'_1 \subset Z \subset Q_1$ et comme Q'_1 et Q_1 ne diffèrent que par q , contenu dans Z , $Z = Q_1$. La partie Q_1 de M satisfait donc aux conditions exigées du reste $R(q)$ dans la définition du bon ordre. C'est d'ailleurs la seule : tout système de parties de M possédant les mêmes propriétés contiendrait M et formerait une θ -chaîne par rapport à φ , donc serait contenu dans M .

L'élargissement par rapport aux notions de *Was sind und was sollen die Zahlen* ? est marqué dans le passage de la chaîne à la θ -chaîne : l'application seule ne permet pas d'atteindre dans l'ordre plus loin que ω . La chaîne propre fait la somme d'une suite infinie d'opérations : elle ne fonde que l'induction complète ordinaire. Pour avoir l'induction complète transfinie — équivalente au bon ordre — il faut introduire une nouvelle opération, ici l'intersection d'un système d'éléments de M_0 ; grâce à elle est garanti que toute partie P de M possède bien une enveloppe dans M_0 . Ainsi se retrouve donc l'indépendance relative des deux principes générateurs des nombres ordinaux chez CANTOR : à l'adjonction d'une unité correspond l'appli-

cation, au passage à la limite, après lequel on peut à nouveau ajouter une unité, l'intersection. Mais si le raisonnement de la chaîne se trouve ainsi compliqué, aucune notion nouvelle n'est introduite, sauf peut-être celle d'ensemble des parties d'un ensemble, que DEDEKIND n'avait pas expressément isolée. La théorie abstraite permet donc d'unifier dans une organisation systématique toute l'œuvre cantorienne ; c'est à quoi s'emploie ZERMELO en 1907.

* * *

Le travail est imposé par une double urgence : l'insertion du théorème du bon ordre dans la théorie classique, la neutralisation des antinomies récemment découvertes. L'axiome de choix est indispensable dans la théorie classique : ZERMELO montre ⁽¹⁾ notamment que sans lui on ne sait ni fonder la théorie des ensembles finis sur la 1^{re} définition de DEDEKIND (démontrer le théorème que tout ensemble fini peut toujours être ordonné de telle sorte que chacune de ses parties possède un premier et un dernier élément) ⁽²⁾, ni prouver les propositions essentielles suivantes : l'ensemble réunion d'un système S d'ensembles disjoints a une puissance supérieure ou égale à celle de S ; les sommes d'ensembles disjoints équivalents sont équivalentes ; une somme dénombrable d'ensembles finis ou dénombrables est dénombrable ⁽³⁾ ; enfin le produit de

(1) ZERMELO (II), p. 113-114.

(2) En effet, ce théorème n'est autre que le théorème DEDEKIND vu plus haut : tout ensemble fini peut être appliqué de façon biunivoque sur un et un seul segment Z_n , conséquence immédiate du théorème que si on peut appliquer de façon biunivoque tout segment Z_n sur un ensemble S , S est infini ; on a vu que dans la démonstration on prenait comme base une application α_n pour chaque Z_n , donc un choix dans l'ensemble des applications de Z_n sur S (signalé pour la 1^{re} fois par HESSENBERG (I)).

(3) Il faut également supposer une application simultanée de tous les éléments de la somme respectivement les uns sur les autres (donc le choix d'une application pour chaque élément).

puissances ne s'annule que si un des facteurs est nul : énoncé équivalent à l'axiome de choix, comme on voit immédiatement ⁽¹⁾. Ainsi avec l'axiome de choix (et le bon ordre) il faut savoir que la théorie à peu près tout entière subsiste ou s'écroule. Mais l'écroulement menaçait déjà par ailleurs.

L'antinomie BURALI-FORTI gênait le libre essor de la théorie ordinaire cantorienne. En 1903, paraît dans la théorie cardinale l'antinomie RUSSELL ⁽²⁾, à laquelle ZERMELO s'était déjà heurté ⁽³⁾. Ici encore c'est un procédé essentiel de la théorie qui, défini abstraitement, procure une contradiction : pour l'antinomie BURALI-FORTI le 2^e principe générateur des ordinaux, pour l'antinomie RUSSELL, le procédé de la diagonale. A toute application d'un ensemble M sur l'ensemble $\mathfrak{P}M$ de ses parties correspond un élément de $\mathfrak{P}M$ extérieur à l'application : c'est l'ensemble des éléments de M qui n'appartiennent pas à la partie à laquelle ils sont coordonnés. Si M est l'ensemble de tous les ensembles, on devrait pourtant avoir $\mathfrak{P}M \subseteq M$: contradiction. La situation est la même dans les deux cas : une théorie s'est constituée pour répondre à des problèmes effectivement posés par la mathématique existante ; les notions, les procédés nouveaux forgés permettent d'obtenir des résultats indépendants de la structure particulière des objets mathématiques dont on parlait. Une théorie abstraite s'impose, dont le seul élément est l'ensemble, « réunion en un tout d'objets déterminés et distincts de notre

(1) Un produit d'ensembles $M, N, P \dots$ a pour éléments les systèmes $\{m, n, p \dots\}$ composés d'éléments appartenant respectivement à $M, N, P \dots$: chaque système représente donc un choix.

(2) RUSSELL-WHITEHEAD (I), p. 363-368.

(3) « J'avais trouvé moi-même cette antinomie indépendamment de RUSSELL et l'avais déjà communiquée... au professeur HILBERT avant 1903. » ZERMELO (II), p. 119 n.

intuition ou de notre pensée » (1). Dès que l'on dit « tous les objets qui... », on a un ensemble. Mais l'ensemble de tous les ordinaux est contradictoire ; mais l'ensemble de tous les ensembles est contradictoire. La faute n'en est pas aux procédés, simple pouvoir exécutif des notions, ce sont les notions bases mêmes qui disparaissent. Qu'est-ce maintenant qu'un ensemble ? et, si l'on change la définition, que reste-t-il de la théorie naïve ? On voit la chute après les espoirs que celle-ci avait fait naître, la gravité pour le système entier des mathématiques, l'impossibilité d'un recours étranger, puisqu'elle est préalable à tout : « La théorie des ensembles, commence ZERMELO, est la branche des mathématiques à laquelle il revient d'étudier mathématiquement les concepts fondamentaux de nombre, d'ordre et de fonction dans leur simplicité primitive, et par là, de développer les bases logiques de l'arithmétique tout entière et de l'analyse (2). »

La seule solution doit être cherchée dans un examen critique : au lieu de poser une nouvelle définition *a priori*, on obtiendra une définition implicite en scrutant raisonnements et résultats qui tiennent : « partant de la « théorie des ensembles » historiquement existante, il faut chercher les principes qui sont nécessaires pour la fondation de cette discipline mathématique . . J'ai l'intention de montrer comment toute la théorie créée par CANTOR et DEDEKIND se laisse ramener à quelques définitions et à 7 axiomes qui semblent indépendants les uns des autres » (3). Travail à la fois d'explicitation (4) et de limitation : il s'agit de « res-

(1) Définition de CANTOR (XXVII), p. 282.

(2) ZERMELO (III), p. 261.

(3) *Ibid.*

(4) ZERMELO l'avait déjà ébauché en tête de sa seconde démonstration du bon ordre, pour prévenir les objections ; outre l'axiome de choix se trouvent énoncés deux postulats qui garantissent l'existence des parties d'ensembles, de l'ensemble des parties d'un ensemble, du complémentaire et de l'intersection d'un système d'ensembles (ZERMELO (II), p. 107).

treindre assez les principes pour exclure toutes les contradictions et ... garder pourtant ce qui a une valeur ». La méthode est celle dont HILBERT avait donné le modèle pour la géométrie : entre les objets indéterminés qui constituent le champ de la théorie (corrélatif des axiomes et non à son tour objet de la théorie, c'est-à-dire non ensemble) l'axiomatique définit par ses propriétés une relation fondamentale ϵ : élément-ensemble ⁽¹⁾. Les seconds termes de cette relation seront les ensembles : ou plutôt la théorie ne s'occupe que de déduire logiquement de nouvelles propriétés à partir des propriétés initiales de la relation, sans s'inquiéter de la réalité ou de la structure des termes. En fait un seul axiome, l'axiome I — de *détermination* — définit une relation : la relation d'égalité suite des relations ϵ et $\subset$ ($M = N$ si $M \subset N$ et $N \subset M$, ou si pour tout $x \in M$, $x \in N$ et réciproquement) tous les autres axiomes sont des affirmations d'existence. La limitation est donc obtenue de façon réaliste : sont ensembles les seuls objets posés comme tels par les axiomes.

Deux sortes d'affirmations : 1^o catégoriques : l'axiome II — des *ensembles élémentaires* — pose comme ensembles l'ensemble vide (simple moyen de supprimer des restrictions dans les raisonnements, mais non indispensable), les ensembles constitués de un et de deux éléments ⁽²⁾ ; l'axiome VII — *de l'infini* — : il existe « un ensemble Z

(1) La relation $M \subset N$ (M *partie* de N) est une abréviation pour : de $x \in M$ suit toujours $x \in N$.

(2) Il s'agit d'affirmation catégorique relativement à l'existence d'ensembles : a étant une chose quelconque du domaine — pas forcément un ensemble — dont l'existence en tant qu'ensemble n'a donc pas à être supposée, $\{a\}$ est un ensemble ; pour $\{a, b\}$, il y a cette fois l'hypothèse que a et b sont distincts, c'est-à-dire que la question $a \in \{b\}$ peut être décidée au moyen des axiomes. Seule l'existence de l'ensemble vide est une affirmation absolument catégorique (v. ci-après son importance dans l'axiomatique de FRAENKEL).

qui contient comme éléments l'ensemble vide et ... avec tout élément a l'élément de forme $\{a\}$ ⁽¹⁾ » ;

2^o Hypothétiques : si un ensemble existe, existent également d'autres ensembles définissables à partir du premier. Axiome III — de *séparation* — : « si une proposition $\mathcal{E}(x)$ est *définie* pour tous les éléments d'un ensemble M , M possède toujours une partie $M\mathcal{E}$ constituée par tous les éléments x de M pour lesquels $\mathcal{E}(x)$ est vrai ⁽²⁾ ». Une proposition est *définie* si « sa validité... peut être décidée sans arbitraire au moyen des axiomes et des lois logiques universelles ». De l'axiome III résultent l'existence de parties ⁽³⁾ d'un ensemble M , de son complémentaire par rapport à un ensemble N qui le contient, de l'intersection d'un système d'ensembles. — L'axiome IV et l'axiome V posent respectivement l'existence de l'*ensemble des parties* $\mathfrak{P}M$ d'un ensemble donné et de l'*ensemble réunion* $\mathfrak{S}M$ formé par les éléments des éléments de M . Enfin l'axiome de *choix*, VI, garantit que le produit n'est pas vide lorsqu'aucun des facteurs n'est vide ; « si T est un ensemble dont tous les éléments sont des ensembles disjoints non vides, leur réunion $\mathfrak{S}T$ contient au moins une partie S_1 qui a, avec chaque élément de T , un et un seul élément commun ⁽⁴⁾ ». Le choix général, dans le cas où les ensembles éléments de T ne sont pas disjoints, s'en déduit, car les axiomes précédents permettent de remplacer toujours un ensemble quelconque T ayant pour éléments des ensembles $A, B, C, \dots$ par un ensemble équivalent T' , dont les éléments $A', B', C', \dots$ sont disjoints et respectivement équivalents à $A, B, C, \dots$

A part l'axiome de choix et l'axiome de l'ensemble des parties, ces énoncés reproduisent exactement — aux

(1) ZERMELO (III), p. 266.

(2) *Ibid.*, p. 263.

(3) A l'exception des parties comme $\{a\}$, $\{a, b\}$ — a et b étant des éléments de l'ensemble — dont l'existence est posée par l'axiome II.

(4) *Ibid.*, p. 266.

termes près — les définitions placées par DEDEKIND dans le § 1 de son ouvrage. L'axiome de l'infini remplace le théorème de l'infini dont la démonstration utilise l'ensemble illégitime de toutes les pensées. On est donc certain de retrouver la théorie classique. ZERMELO en donne la preuve pour la théorie de l'équivalence. Si les démonstrations sont légèrement compliquées, aucun procédé nouveau n'apparaît : le raisonnement de la diagonale et celui de la chaîne se laissent traduire. D'autre part l'antinomie RUSSELL disparaît : de l'axiome III, de *séparation*, résulte que tout ensemble M contient au moins une partie qui n'est pas élément de M ; c'est le procédé de la diagonale qui la donne : soit M la partie formée de tous les éléments qui ne sont pas éléments d'eux-mêmes (propriété *définie*, donc M_0 existe) ; M_0 ne peut être élément de M , sous peine de se contenir comme élément, ce qui serait contradictoire. Or dans l'ensemble de tous les ensembles, s'il existait, toute partie serait en même temps élément. D'une façon plus générale les axiomes ne permettent de poser comme ensemble le système de tous les objets possédant une propriété *définie*, que si les objets sont éléments d'un ensemble déjà reconnu comme tel : ni l'ensemble de tous les ensembles, ni l'ensemble de tous les ordinaux ne peuvent satisfaire à la condition.

Cependant, la théorie des ensembles n'est pas fondée pour cela : en effet, l'œuvre de ZERMELO reste inachevée ; il n'est montré ni que la théorie de l'ordre et des nombres ordinaux peut se déduire des axiomes, ni que ceux-ci sont indépendants entre eux et non contradictoires. D'autre part — et ceci en liaison avec l'absence de démonstration de non contradiction — la forme même des axiomes prête à critique : l'axiome de *séparation* fait intervenir la possibilité d'une décision d'après « les axiomes posés et les lois de la logique ». Or les antinomies de la théorie des ensembles en éveillent d'autres dans le domaine de l'épistémolo-

gie générale, comme celle de RICHARD sur les définitions en termes finis ⁽¹⁾. Toutes les difficultés du problème de la définition mathématique et de celui de la décision se trouvent là, malencontreusement greffées. A combler cette double lacune en modifiant les axiomes et précisant leurs rapports entre eux et avec l'édifice cantorien, s'est attaché FRAENKEL dans une série de mémoires close récemment ⁽²⁾. Trois résultats principaux peuvent en être dégagés :

1° Transformations des notions fondamentales de l'axiomatique — et par suite de l'énoncé même des axiomes ;
 α) transformation du matériel de la théorie : la notion d'un champ général des objets de la théorie, parmi lesquels les ensembles devraient être distingués, est abandonnée ; on exclut tous les objets non-ensembles (à l'exception de l'ensemble vide). Seuls les ensembles sont considérés, et, d'abord — pour la théorie générale, ou théories de l'équivalence, des ensembles ordonnés et bien ordonnés — de façon hypothétique. Il n'y a pour cette théorie aucun axiome d'existence inconditionnée. L'axiome de *détermination* devient : « si a , b et A sont des ensembles et si $a = b$ et $a \in A$ on a aussi $b \in A$ » (l'égalité entre deux ensembles a et b est définie par la relation : pour tout ensemble x , $x \in a$ entraîne $x \in b$ et réciproquement). L'axiome de détermination garantit à cette relation les propriétés de l'égalité mathématique ordinaire : substituabilité de deux objets égaux dans un énoncé de la théorie. L'axiome des *ensembles élémentaires*

(1) Si l'on considère l'ensemble des fractions décimales définissables par un nombre fini de mots : on peut les ranger en une suite dénombrable bien ordonnée et le procédé de la diagonale procure aussitôt une nouvelle fraction extérieure à la suite et pourtant définissable également avec un nombre fini de mots : le procédé de la diagonale n'est pas responsable de la contradiction, mais l'incertitude de la « définition en termes finis ». Ces antinomies « épistémologiques » ne concernent pas directement la théorie des ensembles.

(2) V. en particulier FRAENKEL (VII) et (X) où les résultats de ces mémoires sont résumés.

taires se réduit à un axiome d'*accouplement* (affirmation d'existence hypothétique) : si a et b sont des ensembles distincts, il existe un ensemble ayant pour éléments a et b : $\{a, b\}$;

β) transformation principale : la notion de fonction définie inductivement remplace la notion de propriété *définie*. L'axiome de *separation* s'énonce : s'il existe un ensemble M et deux fonctions données $\varphi(x)$ et $\psi(x)$, il existe une partie de M comprenant tous les éléments x de M pour lesquels $\varphi(x) \in \psi(x)$ ou $\varphi(x) \notin \psi(x)$. Comme fonctions de l'ensemble x sont posées : d'abord les constantes (c'est-à-dire les ensembles admis comme existant inconditionnellement), x lui-même, l'ensemble des parties de x ($\mathfrak{P}x$), l'ensemble réunion $\mathfrak{E}x$, l'ensemble couple $\{\varphi(x), \psi(x)\}$ (si $\varphi(x)$ et $\psi(x)$ sont déjà connus) enfin l'ensemble fonction de fonction $\varphi(\psi(x))$. On le voit, le progrès sur ZERMELO a consisté à interpréter la phrase : « propriété décidable au moyen des axiomes de la théorie », en termes mêmes de la théorie (au moyen de la relation fondamentale ϵ et des fonctions que les axiomes II, IV, V de ZERMELO procurent ; l'axiome VI ne définissant pas de façon unique, aucun ensemble de choix ne peut intervenir). Mais le procès inductif ne s'arrête évidemment pas là : si dans la définition de l'ensemble M ou des fonctions $\varphi(x)$, $\psi(x)$ intervient une autre variable y , la partie d'ensemble :

$$M_{\varphi(x) \in \psi(x)}$$

dépend aussi de y , elle constituera donc aussi une fonction de y . A la vérité une difficulté apparaît dès le début du procès si, par exemple, pour la définition d'une partie de M l'ensemble $\mathfrak{P}M$ des parties de M est utilisé : c'est là, note FRAENKEL, « un exemple classique de définition non prédicative » (1). La solution est celle qu'admettent les mathé-

(1) FRAENKEL (VII), p. 112.

maticiens non intuitionnistes en pareil cas : une construction de la partie en question est impossible, mais il peut y avoir un sens mathématique dans la caractérisation d'un élément d'un ensemble, fût-ce au moyen de la totalité de l'ensemble (par exemple la borne supérieure ou inférieure en analyse).

2° Sectionnement de la théorie en fragments rattachés à un nombre croissant d'axiomes. Grâce aux axiomes précédents auxquels sont joints sans modification les axiomes de l'ensemble des parties, de l'ensemble réunion, et du choix (IV, V, VI de ZERMELO) peut être construite la théorie générale des ensembles ⁽¹⁾ : théorie de l'équivalence, théorie des ensembles ordonnés et bien ordonnés dont FRAENKEL a opéré le développement. En particulier sont assurées, au cas où en général un ensemble existe (ce qui n'a pas encore été posé), l'existence de l'ensemble vide et celle de l'ensemble formé d'un élément ⁽²⁾. L'ordre est défini par le procédé de subsomption des restes (introduit par HESSENBERG ⁽³⁾) : un ensemble M est ordonné s'il existe une partie de $\mathfrak{C}$ de l'ensemble des parties $\mathfrak{P}M$ de M , telle que : 1° de deux éléments de $\mathfrak{C}$ l'un est partie de l'autre et 2° $\mathfrak{C}$ est une partie saturée de $\mathfrak{P}M$ possédant cette propriété (c'est-à-dire la relation $\mathfrak{C} \subset \mathfrak{C}'$, $\mathfrak{C}'$ possédant la propriété 1° entraîne $\mathfrak{C} = \mathfrak{C}'$) ⁽⁴⁾. Avec l'axiome de l'infini sous la forme

(1) Note de FRAENKEL in VON NEUMANN (IV), p. 392.

(2) FRAENKEL (V), p. 256.

(3) HESSENBERG (I) ; cf. aussi FRAENKEL (X) et (VII), p. 137 et KURATOWSKI (I).

(4) On dit que $\mathfrak{C}$ ordonne M : la 1^{re} propriété établit un ordre entre tous les éléments de M tels que deux quelconques d'entre eux a et b appartiennent respectivement à deux éléments distincts $T(a)$ et $T(b)$ de $\mathfrak{C}$; on a :

$$a \prec b \text{ si } T(a) \supset T(b)$$

(une des 2 relations entre $T(a)$ et $T(b)$ existe toujours ; relation transitive) ; la 2^e propriété garantit que l'ordination est complète (étendue à tous les éléments de M).

donnée par ZERMELO sont posées les existences classiques : ensembles élémentaires, dénombrable, continu et ensembles dont les puissances sont de forme $2^{2^{n_0}}$, $2^{2^{2^{n_0}}}$, etc... .. En revanche, l'existence de l'ensemble dénombrable

$$\{Z, \mathfrak{P}Z, \mathfrak{P}\mathfrak{P}Z, \mathfrak{P}\mathfrak{P}\mathfrak{P}Z, \dots\}$$

où Z représente la suite des entiers, ne peut être assurée, non plus que la théorie des ordinaux transfinis, ce qui était déjà le cas dans l'axiomatique de ZERMELO. FRAENKEL énonce dans ce but un axiome VIII de *substitution* ⁽¹⁾ : « Si M est un ensemble et $\varphi(x)$ une fonction, il existe un ensemble $\mathfrak{M}$ déduit de M par substitution de $\varphi(x)$ à chacun de ses éléments x ». Il n'y a gain, ainsi que l'a montré VON NEUMANN ⁽²⁾, que si on élargit la notion de fonction : $\varphi_y(x)$ et $\psi_y(x)$ étant deux fonctions de x dépendant du paramètre y , « si pour tout y (parcourant un certain ensemble) existe l'ensemble $\Phi(y)$ des x tels que :

$$\varphi_y(x) \in \psi_y(x),$$

$\Phi(y)$ est une fonction de y ». L'élargissement vient de ce que les ensembles des x , $\Phi(y)$, ne sont plus — comme dans les précédentes définitions — astreints à être parties d'un certain ensemble majorant T , déjà posé, le même pour tous les y , c'est-à-dire tel qu'on aurait, pour chaque y :

$$\Phi(y) = T_{\varphi_y(x)} \in \psi_y(x)$$

On obtient donc davantage de fonctions ⁽³⁾. Telles sont les trois principales étapes dans le développement axiomatique de la théorie cantorienne. Doit-il y en avoir d'autres, si l'on

(1) FRAENKEL (X), p. 309.

(2) VON NEUMANN (IV), p. 377 : avec les fonctions définies comme précédemment VON NEUMANN montre que l'axiome de substitution est une conséquence des autres axiomes.

(3) VON NEUMANN (IV), p. 376.

veut retrouver tous les résultats que la continuation des méthodes naïves ⁽¹⁾ a permis d'atteindre après CANTOR ? la question est encore en suspens.

3^o Démonstration d'indépendance relative entre les axiomes ⁽²⁾ : il n'y a là, semble-t-il, qu'exigence esthétique pour l'édification de la théorie. Aussi bien se trouve-t-elle satisfaite en gros par la façon même dont ils sont posés les uns après les autres comme garantie indispensable pour l'existence (hypothétique ou catégorique) de certains ensembles. Seule l'indépendance de l'axiome de choix concerne directement le contenu de la théorie classique : les objections contre la démonstration de ZERMELO venaient, on l'a vu, de la faible distance entre hypothèse et conclusion. Au lieu de s'attaquer à la validité de l'hypothèse — donc de rejeter simultanément les deux propositions, puisque équivalentes — une autre voie serait de la supprimer en tant qu'hypothèse, parce que démontrable à partir des axiomes déjà posés. Le particulier de son énoncé même y invite : puisqu'il garantit, étant donné un ensemble M d'ensembles, l'existence d'une certaine partie de l'ensemble réunion $\mathfrak{S}M$; il devrait tomber sous la juridiction de l'axiome de séparation (par quoi sont posées en général les existences de partie d'ensembles) : en progressant assez dans la construction des fonctions introduites par les définitions de FRAENKEL, pourrait-on atteindre un ensemble de choix ? FRAENKEL a réussi sous certaines conditions à montrer qu'il n'en est rien. Il distingue quatre énoncés, de forces en gros croissantes : VI_1 existence d'un ensemble de choix pour un ensemble dénombrable d'ensembles finis ; VI_2 pour un ensemble

(1) Cf. les travaux de MAHLO (I) et (II), H. HAUSDORFF (I) et BAER (I) sur les nombres initiaux réguliers à indices limites, cités par FRAENKEL (X), p. 310.

(2) Les trois étapes précédentes n'indiquent nullement un ordre essentiel entre les axiomes : chaque axiome est indépendant des précédents et des suivants.

dénombrable d'ensembles quelconques ; VI_3 pour un ensemble quelconque d'ensembles finis ; VI_4 pour un ensemble quelconque d'ensembles quelconques. L'indépendance pour eux se définit ainsi : VI_1 peut n'être pas satisfait lorsque tous les axiomes $I-V$ et VII de ZERMELO-FRAENKEL sont satisfaits ; VI_2 , VI_4 peuvent chacun n'être pas satisfaits lorsque $I-V$, VI_1 (éventuellement aussi VI_3), et VII sont satisfaits. La démonstration a été effectuée par FRAENKEL pour un domaine différent de celui que considère son axiomatique. Au lieu de se restreindre aux seuls objets ensembles, il considère — comme faisait primitivement ZERMELO — d'autres objets de la théorie, qui ne sont pas des ensembles, et à partir desquels l'axiome d'accouplement en particulier permet d'en former. La définition de l'égalité (pour ces objets elle dit exiger l'identité) doit être modifiée en conséquence. En revanche les autres amendements FRAENKEL (nouvelle forme de l'axiome de séparation) sont maintenus et jouent un rôle essentiel dans la démonstration : on procède par induction complète ⁽¹⁾ sur la construction de tous les ensembles qui peuvent être obtenus dans la théorie à partir d'une certaine base définie ; la base et tous les ensembles

(1) On construit une pseudo-théorie où figure, pour le cas VI_1 un ensemble dénombrable de couples d'objets qui ne sont pas eux-mêmes des ensembles :

$$P = \{(a_1 b_1), (a_2 b_2), \dots\}$$

et on démontre que tout ensemble M de la théorie est invariant si on permute entre eux les éléments accouplés dans une certaine partie P_m de P comprenant presque tous les couples de P (c'est-à-dire tous sauf un nombre fini). Ainsi l'ensemble $\{a\}$ est invariant par rapport aux permutations effectuées dans les couples de l'ensemble comprenant tous les couples de P , sauf $(a b)$. Donc il ne peut exister pour P d'ensemble de choix (ayant un élément et un seul commun avec chacun des couples) : il ne serait invariant par rapport à aucune partie de P . Pour les axiomes VI_2 , VI_4 on remplace dans P les couples par des ensembles quelconques, pour les axiomes VI_3 et VI_4 , P n'est pas assujéti à être dénombrable. Cf. FRAENKEL (VII), p. 165 sq. et (XIII), p. 42-51. La démonstration pour VI_1 a été publiée in FRAENKEL (III). Dans FRAENKEL (XIV) vient d'être donnée la démonstration détaillée pour l'indépendance de l'axiome général (VI_4) par rapport au système des axiomes $I-V$, VII et VI_1 .

qui en dérivent possèdent une propriété qui exclut l'existence de l'ensemble de choix pour certains ensembles d'ensembles. Pour un domaine où apparaissent les seuls ensembles aucune démonstration n'est en vue : preuve que cette élimination des objets non ensembles n'est pas inoffensive : elle semble nécessaire si l'on veut préciser le caractère mathématique de la théorie.

A cet achèvement de l'œuvre zermelienne, manquait encore une définition des nombres ordinaux transfinis. Les travaux de Von NEUMANN ⁽¹⁾ à ce sujet, outre leur intérêt intrinsèque, permettent de fonder les définitions par récurrence transfinie et — par le détour d'une autre axiomatique — ouvrent de nouvelles perspectives sur les rapports entre elles des différentes parties de la théorie cantorienne. Le nombre ordinal peut être caractérisé (de façon traduisible dans le système ZERMELO-FRAENKEL) comme un ensemble bien ordonné dont chaque élément est égal au segment qu'il détermine dans l'ensemble. On retrouve le procédé naïf de la numération : compter les éléments d'un système est bien remplacer chaque système partiel déjà compté par un nombre, à la fois élément succédant aux nombres précédents, et représentant de leur collectivité. L'ordre d'un nombre ordinal, ou d'un système de nombres ordinaux est l'ordre de subsomption, introduit par HESSENBERG ⁽²⁾ : les relations $P < Q$ et $P \subset Q$ sont équivalentes (immédiat d'après la définition). A tout ensemble bien ordonné correspond un nombre ordinal et un seul, le nombre ordinal qui lui est semblable. Par induction complète transfinie : d'une part deux nombres ordinaux semblables sont identiques, dans l'application de l'un sur l'autre le premier élément qui serait distinct de son image est égal au segment qui le pré-

(1) VON NEUMANN (I), (III) et (IV).

(2) HESSENBERG (II).

cède, segment par hypothèse identique à son image. D'autre part la propriété pour un ensemble bien ordonné d'être semblable à un nombre ordinal se transmet de l'ensemble à ce segment (par suite de l'identité entre un segment d'un segment et un segment de l'ensemble) et des segments à l'ensemble. Pour la deuxième transmission — qui reprend le procédé de DEDEKIND pour appliquer la suite des entiers sur un ensemble lorsque tout segment de cette suite y est applicable — l'axiome de substitution (non ici l'axiome de choix puisque l'application est déterminée de façon univoque, un ensemble bien ordonné ne pouvant être semblable qu'à un seul nombre ordinal) doit être utilisé : des fonctions bien déterminées appliquant chaque segment de M sur un nombre ordinal, se déduit une fonction $f(x)$ faisant correspondre à l'élément x , définissant un segment, le nombre ordinal sur lequel il est appliqué. Les valeurs de $f(x)$ constituent, en vertu de l'axiome de substitution, un ensemble dont on voit sans peine qu'il possède les propriétés d'un nombre ordinal et qu'il est semblable à M . L'induction complète transfinie achève le raisonnement : la *numérabilité* ⁽¹⁾ (similitude avec un nombre ordinal) se transmet le long du bon ordre.

Deux propriétés jouent surtout un rôle dans la théorie : pour aucun ensemble qui soit nombre ordinal, x , on ne peut avoir $x \in x$ (conséquence immédiate de la théorie du bon ordre) ; on peut prendre comme propriété caractéristique du nombre ordinal : ensemble P de nombres ordinaux tel que tout élément soit en même temps une partie (si $x \in P$, $x \subset P$). Les procédés d'engendrement de CANTOR — dont on a vu qu'ils étaient englobés dans la définition du bon ordre — peuvent alors être déduits. Si M est un ensemble de nombres ordinaux, $\mathfrak{S}M$ sera un nombre ordinal (en vertu

(1) En allemand : *Zählbarkeit* ; à distinguer de la *dénombrabilité* (*Abzählbarkeit*).

de la définition précédente). Le passage à la limite, après une suite N bien ordonnée de nombres ordinaux, consiste alors à considérer l'ensemble réunion $\mathfrak{E}N$; ce nouveau nombre ordinal suit bien immédiatement tous les éléments de la suite (si elle n'a pas de dernier élément) : on sait en effet que si x est un nombre ordinal et :

$$x \prec \mathfrak{E}N \quad \text{on a } x \subset \mathfrak{E}N \text{ donc } x \in N.$$

L'adjonction d'une unité se fait par considération de l'ensemble x qui est un nombre ordinal (également en vertu de la définition précédente). On démontre aisément que pour tout nombre ordinal existe l'alternative :

$$x = \mathfrak{E}x \text{ (nombre de 2^e espèce) ou } x = \mathfrak{E}x + \{ \mathfrak{E}x \}$$

(nombres de 1^{re} espèce) : on voit que si l'ensemble des nombres précédant x a un dernier élément y , on a $y = \mathfrak{E}x$; on peut donc écrire :

$$x = y + 1$$

La théorie des définitions par récurrence transfinie ⁽¹⁾ se déroule dans une symétrie parfaite avec celle de DEDEKIND pour les définitions par récurrence finie. Il s'agit de prouver l'existence d'une fonction et d'une seule $f(x)$ définie pour tout nombre ordinal x et satisfaisant à :

$$f(x) = \varphi(F(f, x), x) \quad (1)$$

φ étant une fonction donnée et $F(f, x)$ représentant l'ensemble des valeurs de f pour les éléments du nombre ordinal x . On démontre par induction complète transfinie que pour tout nombre ordinal y existe un *élément de fonction* $f_y(x)$ et un seul, défini pour tout $x \prec y$ et satisfaisant pour ces valeurs à (1) : la propriété se transmet en effet d'un ensemble bien ordonné à tous ses segments et de ceux-ci à l'ensemble. Si l'on fusionne en posant :

(1) Le procédé de définition par récurrence transfinie tel que l'employaient HAUSDORFF et FRAENKEL n'était pas encore rigoureusement fondé.

$$f(x) = f_x(x)$$

on a bien la fonction cherchée.

Les résultats précédents ne dépendent que de la théorie du bon ordre, sans théorème du bon ordre. Ils ne font intervenir ni l'axiome de choix ni la théorie de l'équivalence, au moins dans sa totalité. D'où l'idée d'inverser le développement traditionnel du système cantorien, idée que CANTOR semble avoir eue lui-même. L'axiomatique établie par VON NEUMANN se distingue en outre d'une part en considérant comme objets des fonctions et des arguments, au lieu d'ensembles et d'éléments (différence technique accessoire, destinée seulement à simplifier) ; d'autre part en ne limitant pas le champ de la théorie par exclusion des objets « trop gros », sources des paradoxes connus, mais en imposant seulement que toutes les fonctions ne puissent devenir arguments. Les contradictions n'apparaissent en effet que lorsque l'ensemble de tous les ensembles ou de tous les ordinaux sont traités en éléments possibles d'ensembles. Tous les arguments peuvent être fonctions (simplification ⁽¹⁾) analogue à celle de FRAENKEL par rapport au système ZERMELO), mais la condition nécessaire et suffisante pour qu'une fonction ne puisse être argument sera qu'à tout argument du domaine corresponde de façon bi-univoque un argument pour lequel cette fonction a une valeur différente d'un argument A donné une fois pour toutes. Si l'on appelle *élément x du système S* , tout argument pour lequel la fonction :

$$f(x) \neq A,$$

on voit que pourront être éléments et eux seuls, les systèmes ne pouvant être appliqués de façon bi-univoque sur le système de tous les arguments (*Axiome N*) : les systèmes pouvant être éléments sont les ensembles.

(1) Simplification que ne fait d'ailleurs VON NEUMANN ni dans (II) ni dans (III) mais seulement dans (V).

Cet axiome remplace à la fois, dans l'axiomatique ZERMELO-FRAENKEL, les axiomes d'accouplement, de séparation, de substitution et de choix. Il écarte évidemment l'antinomie RUSSELL : le système de tous les ensembles n'est pas un ensemble. Quant à l'antinomie BURALI-FORTI, elle se transforme en *théorème du bon ordre*, suivant le procédé que CANTOR avait essayé avec la distinction entre ensembles consistants et inconsistants ⁽¹⁾. Le système Ω^{**} de tous les ordinaux ne peut être un ensemble : en effet on a vu que tout ensemble d'ordinaux dont les éléments sont à la fois ses parties (condition évidemment remplie) est un nombre ordinal : donc ce nombre ordinal serait son propre élément, ce qu'on a vu également être impossible ⁽²⁾. S'il n'est pas un ensemble, il est, en vertu de l'axiome *N*, applicable sur le système Ω de tous les éléments (ou arguments). Cette application bi-univoque communique à Ω le bon ordre de Ω^{**} : comme tout ensemble est une vraie partie de Ω , il est bien ordonnable.

Le raisonnement a l'avantage de mettre nettement en lumière le caractère spécial de l'antinomie : c'est une propriété particulière aux nombres ordinaux (découlant de la théorie du bon ordre : un segment d'un ensemble bien ordonné ne peut être semblable à un segment qui en est une vraie partie) qui empêche Ω^{**} d'être un ensemble. En revanche, la conclusion est entachée de tout l'arbitraire de l'axiome *N* : celui-ci pose l'existence d'une fonction qui applique le système non ensemble (ici Ω^{**}) sur le système Ω de tous les arguments ; d'elle se déduit la fonction φ bien ordonnante de toute partie de Ω , et également la

(1) V. ci-dessus, p. 95.

(2) Dans (III) VON NEUMANN admet qu'un système (non ensemble) puisse être un nombre ordinal : mais tous ses segments doivent être des ensembles puisqu'ils figurent comme argument dans la fonction faisant correspondre au segment l'élément qui le détermine. Ω^{**} est alors le seul nombre ordinal qui ne soit pas un ensemble : rien n'est changé au raisonnement du texte.

fonction ψ de choix qui distingue, pour chaque ensemble, le premier élément dans le bon ordre imposé par φ . La théorie de l'équivalence se peut ensuite facilement développer : on appelle nombre cardinal un nombre ordinal qui ne peut être appliqué de façon bi-univoque sur aucune de ses vraies parties ; l'universelle comparabilité des puissances se déduit de celle des nombres ordinaux.

On retrouve donc en pleine lumière ce qui n'était que latent chez CANTOR : il n'y a, en réalité, qu'une seule catégorie de nombres, les nombres ordinaux ; le caractère numérique est lié aux propriétés de l'acte de compter d'où sortent directement bon ordre et nombres ordinaux. Ce n'est que l'intervention de considérations étrangères (de nouvelles fonctions), qui parmi les ordinaux en distingue certains doués d'une propriété particulière, les nombres initiaux de classe. Mais ici encore apparaît un accident dans la création : le troisième principe d'arrêt. On peut faire toute la théorie des ordinaux sans y avoir recours, développer toute l'arithmétique ordinale transfinie non seulement sans lui, mais sans même s'inquiéter d'une pluralité de puissances.

C'est, en réalité, la mathématique classique qui exige leur affirmation. Mais ici apparaît une dernière difficulté que Von NEUMANN a contribué à préciser ; les comparaisons de puissances — et d'une façon générale les propriétés de bon ordre, etc. — sont relatives à l'axiomatique considéré. Si, à un système Σ_1 satisfaisant aux axiomes de Von NEUMANN, on superpose un système Σ_2 , soumis pour l'essentiel aux mêmes axiomes, mais comportant davantage de fonctions (ce qu'il est possible d'obtenir par des modifications accessoires des axiomes), le même objet a pourra être dans Σ_1 un ensemble non dénombrable et se trouver dénombrable dans Σ_2 ; ainsi également pour la finitude et le bon ordre : « il n'est pas exclu que l'ensemble a paraisse fini et bien ordonné dans le système Σ_1 et se révèle dans le système « plus fin » Σ_2

non bien ordonné (infini) simplement parce qu'une partie de a , qui n'a pas de premier élément, n'était pas un ensemble dans Σ_1 , donc n'avait pas été remarqué, tandis qu'elle est un ensemble dans Σ_2 » (1). Donc relativité aussi bien vers le bas (finitude et bon ordre) que vers le haut.

Même relativité dévoilée antérieurement par le paradoxe de SKOLEM : chaque système d'axiomes est satisfait par une infinité dénombrable d'objets (par exemple des ensembles que procurent les axiomes ZERMELO-FRAENKEL) (2). Il s'agit évidemment d'une autre dénombrabilité que de celle qui est définissable au sein de l'axiomatique. Mais le divorce même est significatif : entre le mouvement naïf des créations mathématiques et les systèmes axiomatisés la jonction immédiate est impossible. On a l'impression de microcosmes, ingénieuses miniatures « qui représentent fidèlement la théorie des ensembles jusque dans tous les détails.. Il y paraît toutes les puissances... Mais aussitôt que l'on met en jeu des instruments plus fins tout s'évanouit. De toutes les puissances ne restent que le fini et le dénombrable. Cela seul a un sens réel, tout le reste est fiction formaliste (3) ». Au moins l'utilisation dans des problèmes mathématiques en est compliquée.

L'axiomatisation de la théorie des ensembles n'était, dans l'intention, que codification de démarches effectives pour l'élimination des antinomies. On ne peut guère prouver que le résultat soit atteint : si les ensembles de RUSSELL et de BURALI-FORTI disparaissent, MIRIMANOFF a donné l'exem-

(1) VON NEUMANN (II), p. 240.

(2) La démonstration rigoureuse ne peut s'effectuer que dans un système entièrement formalisé, quoique SKOLEM ait montré que le résultat est indépendant du mode de formalisation ; la construction récurrente d'ensembles, telle qu'elle s'effectue à partir du seul axiome d'existence absolue, axiome de l'infini, dans le système ZERMELO-FRAENKEL, grâce à l'axiome de séparation modifié, en peut donner idée.

(3) VON NEUMANN, *ibid.*, p. 239.

ple ⁽¹⁾ d'un ensemble descendant, impossible à exorciser avec les axiomes connus si l'on ne veut pas se borner aux ensembles construits par les axiomes (c'est-à-dire à une infinité dénombrable) comme cela est possible dans les systèmes FRAENKEL et VON NEUMANN, et sur lequel aucun calcul n'est possible. Seule une démonstration de non contradiction donnerait la sécurité : en 1907 ZERMELO avouait n'y être pas encore arrivé ⁽²⁾. Toutes les démonstrations de non contradiction connues consistent à fabriquer un modèle de la théorie, dans une théorie mathématique déjà fondée : pour la théorie des ensembles, le procédé est évidemment inapplicable. En revanche la difficulté du relativisme enlève aux démarches axiomatisées leur efficacité concrète : si l'on veut garder sa place à la théorie, il faut que, pierre d'angle des mathématiques, elle n'en soit pas séparée dans l'opération d'axiomatisation. C'est à une reconstruction complète de l'édifice qu'il importerait de se livrer, qu'elle aboutisse au moyen d'une axiomatisation rigoureuse, ou négativement — avec l'intuitionnisme — par définition nouvelle des enchaînements admissibles : dans tous les cas le problème de la théorie des ensembles devient problème du fondement des mathématiques.

(1) MIRIMANOFF (I) : il s'agit de l'ensemble $a = \{a\}$ avec $\{a\} = \{\{a\}\}$, etc...

(2) ZERMELO (III). p. 262.

CONCLUSION

Il est possible de donner au procès d'axiomatisation une signification moins absolue : sorte de condensation de la matière active d'une théorie autour de quelques procédés intuitivement saisissables et pour lesquels la question d'une approbation, sinon d'une justification, est plus facile à résoudre. De ce point de vue les recherches de ZERMELO-FRAENKEL et Von NEUMANN sont aussi utiles pour le développement de la théorie, que significatives dans le dessin de son histoire. Grâce à elles semblent se fixer — du moins autant qu'il se peut actuellement — portée du système et liaisons véritables.

La théorie de l'équivalence est née au sein de l'analyse, avec des méthodes empruntées à l'analyse. Si BOLZANO pressentait déjà les notions de la réunion et du produit d'ensembles, WEIERSTRASS, à propos des nombres complexes, en donnait la définition précise que CANTOR n'eut qu'à transposer ⁽¹⁾ : opérations simples, résultats pauvres, tout l'intérêt de la théorie est dans la manipulation qu'elle rend plus aisée, dans les mises en rapport qu'elle établit pour des ensembles déjà procurés par l'analyse. Si on la veut préalable à celle-ci, elle n'existe que dans la mesure où sa liaison avec elle, de contiguïté devient continuité, sous peine de s'évanouir dans le vide logique. Nombre, fonction, ces actes intuitifs de l'esprit sont des instruments qui se forgent eux-mêmes, car rien n'existe avant eux. La duperie de l'application comme opération logique antérieure à la fonction se manifeste dans la floraison des paradoxes : c'est, après ZERMELO, la précise définition de l'opération de séparation ⁽²⁾, par construction récurrente de fonctions, qui

(1) V. ci-dessus, p. 18.

(2) SKOLEM s'y était déjà employé in SKOLEM (II).

permet à FRAENKEL de relever l'édifice de Dedekind. Après les fonctions réunion, ensemble des parties, etc., le moment non prédicatif permet une génération illimitée qui ouvre les perspectives de la théorie. Peut-elle aller très loin avec ses seules ressources (sans axiome de substitution) ? Il est difficile d'en décider : outre la définition de l'ordre et du bon ordre, elle n'a guère obtenu de résultats après les théorèmes de BERNSTEIN et de KOENIG. Pour être fécondée il semble qu'il lui faille des objets — ensembles ou fonctions — définis dans la mathématique classique : fonction additive, mesure dans l'intégration abstraite, notion de voisinage pour les espaces topologiques. Résultats plus nombreux, sécurité gagnée tiennent à la solidarité mathématique intérieure des actes qu'ils réclament, au contrôle effectif auquel les soumet leur caractère plus concret. Il n'y a plus alors développement singulier d'une théorie autonome. Le problème de l'universelle comparabilité est assez révélateur de l'imagination, qui sous-tendait les travaux de CANTOR, d'un en-soi des ensembles : comment affirmer l'existence de la fonction appliquante, si les deux ensembles en présence ne sont définis par rien d'autre que leur existence ? il semble que les opérations de l'équivalence ne permettent pas de l'atteindre. Au lieu de problème il n'y a sans doute là qu'un nouvel axiome, dont HARTOGS ⁽¹⁾ a montré qu'il équivaut à l'affirmation du bon ordre (et par suite à l'axiome de choix) : son indépendance par rapport aux autres axiomes résulte donc des travaux de FRAENKEL sur l'axiome de choix.

Avec la théorie des ordinaux, apparaît un mouvement de pensée radicalement nouveau — caractérisé par l'axiome de substitution — plus propre, semble-t-il, à procurer le développement d'une véritable théorie. On a vu que sa subordination aux considérations d'équivalence résulte d'un accident historique. Mais il faut encore noter que ce n'est pas le bon ordre qui est préalable au nombrement, mais le nombre-

(1) Il est évident que du bon ordre résulte la comparabilité (puisque tous les ensembles bien ordonnés sont comparables). HARTOGS a prouvé la réciproque. HARTOGS (I).

ment au bon ordre ⁽¹⁾. Aucun ensemble mathématique effectif, hors la suite des entiers, ne donne exemple de bon ordre ; s'il est définissable avec les axiomes de l'équivalence, aucun moment intérieur de cette théorie n'y conduirait si l'idée n'en était fournie par l'engendrement des ordinaux. Il semble donc que — quels que soient par ailleurs ses avantages — la présentation axiomatique ne fasse ici que reprendre, à peine masqués, les enchaînements naïfs : la validité de celle-là, en l'absence de toute démonstration de non contradiction, ne pourrait s'appuyer que sur la réalité de ceux-ci ; retour aux principes d'engendrement de CANTOR. Le remarquable est d'ailleurs qu'une axiomatisation comme celle de VON NEUMANN d'une part élimine toute une série de raisonnements de la théorie de l'équivalence (démonstration Bernstein, démonstration Zermelo du bon ordre), d'autre part n'apporte aucune lumière sur le problème du continu : s'il est possible d'y prouver que la puissance d'un ensemble M est inférieure à celle de l'ensemble de ses parties $\mathfrak{P}M$ on n'a aucun moyen de comparer $2^{\aleph_0}$ et $\aleph_1$. Certaines existences posées par l'axiome N — ou par l'axiome de substitution de FRAENKEL — et celles données par d'autres axiomes, comme celui de l'ensemble des parties ou de l'ensemble réunion, semblent impossibles à mettre en rapport autrement que par un nouvel axiome. C'est aussi bien ce que rend vraisemblable la possibilité d'une deuxième hypothèse du continu, reprise par LUSIN ⁽²⁾, et qui, jusqu'ici, n'a donné lieu à aucune contradiction, bien qu'elle exclue la première : $2^{\aleph_0} = 2^{\aleph_1}$.

Sa place privilégiée, les difficultés rencontrées, donnent à l'histoire de la théorie des ensembles sa portée philosophique : l'ambition d'un développement préalable au reste des mathématiques devait permettre de préciser la notion de théorie mathématique, isolée comme *in vitro* ; l'obliga-

(1) D'autre part nous ne connaissons aucun moyen d'ordonner un ensemble autrement que par le bon ordre, tandis que nous connaissons des ensembles simplement ordonnés.

(2) Cf. LUSIN (I).

tion, après les paradoxes, d'énumérer exhaustivement les hypothèses, précise la notion même de démonstration : si les résultats ne sont qu'une transformation immédiate des axiomes ⁽¹⁾, peut-on encore parler de théorie ? La réalité d'un raisonnement se trouve en quelque sorte mesurée par son brassage de matière mathématique, par l'unification qu'il procure entre objets — procédés — déjà mathématiquement existant. D'où la difficulté de divisions de droit dans un tissu vivant, par rapport auquel les affirmations ne sont que superstructures des procédés, où l'imprévisible déviation d'un geste procure à la fois les nouveautés et maintient les liaisons. Rejeter ou fonder une théorie n'est travail définitif, ni dépourvu de degrés ; surtout, comme pour les caractérisations d'unité, il ne peut s'opérer par simple investigation logique. Si les axiomatiques semblent dissocier deux théories : équivalence, nombres ordinaux (et peut-être une troisième où serait étudié le problème du continu), dont la sécurité va décroissant, les considérations pragmatistes du mathématicien militant ont le dernier mot. C'est dans l'analyse entière que se pose le problème d'une existence et d'un affermissement, pour lesquels une définition du travail mathématique en général est nécessaire.

(1) Il ne s'agit ici bien entendu en aucun cas des importantes démonstrations données par ZERMELO, FRAENKEL, VON NEUMANN au cours de leurs axiomatisations respectives. Mais il y a peu de théorèmes.

Vu : le 8 juillet 1937,
Pour le Doyen de la Faculté des Lettres
de l'Université de Paris,

J. VENDRYES.

Vu et permis d'imprimer :
Le Recteur de l'Académie de Paris,
S. CHARLÉTY.

BIBLIOGRAPHIE (1)

OUVRAGES GÉNÉRAUX (2)

(cités dans tout le cours du livre).

- CANTOR G. (XXVIII). — *Gesammelte Abhandlungen*, herausggben v. E. Zermelo, nebst einem Lebenslauf Cantors von A. Fraenkel. Berlin, Springer, 1932.
- DEDEKIND R. (V). — *Gesammelte mathematische Werke*, herausggben v. R. Fricke, E. Noether u. O. Ore. Braunschweig Vieweg u. Sohn, 1931, 3 vol.
- CANTOR-DEDEKIND (I). — *Briefwechsel* herausggben von. E. Noether u. J. Cavallès, Paris, Hermann, 1937.
- FRAENKEL A. (X). — *Einleitung in die Mengenlehre*, 3^e ed. Berlin, Springer, 1928, t. 1, 420 p.
- (XI). — *Georg Cantor*. Jahresber. d. deutsch. Math. Verein., 1930.
- HAUSDORFF F. (I). — *Grundzüge der Mengenlehre*. Leipzig Veit, u. C., 1914.
- (II). — *Mengenlehre* (2^e éd. modifiée du précédent). Berlin W. de Gruyter, 1927.
- HESSENBERG G. (I). — *Grundbegriffe der Mengenlehre* (Sonderdruck aus den *Abhandl. d. Friesschen Schule*). Göttingen, 1906.
- SCHOENFLIES A. (I). — *Die Entwicklung der Lehre von den Punktmannigfaltigkeiten* I. Teil: *Jahresb. d. deutsch. Math. Verein.* t. 8 (1900), II. Teil: 2 Ergänzungsb. d. *Jahresb. d. deutsch. Math. Verein.* (1908). Leipzig.

CHAPITRE I

§ 1-2. — *Le continu et l'introduction des nombres irrationnels.*

- BOLZANO (I). — *Rein analytischer Beweis des Lehrsatzes dass zwischen je zwei Werten die ein entgegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege*. *Abhdl. d. Kgl. Ges.*

1. Il ne s'agit ici que d'ouvrages en liaison directe avec le texte. On trouvera une bibliographie complète sur la théorie des ensembles proprement dite (jusqu'à 1928) in Fraenkel (X).

2. La pagination de tous les mémoires cités de Cantor et de Dedekind est empruntée à celle des deux recueils Cantor (XXVIII) et Dedekind (V).

- der Wiss. Prag.*, 1817. Herausggben v. Ph. E. B. Jourdain (Ostwald's Klassiker). Leipzig Engelmann, 1905.
- CANTOR (V). — *Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen. Math. Annalen*, t. 5 (1872), p. 123-132.
- CANTOR (XV). — *Über unendliche lineare Punktmannigfaltigkeiten* 5. *Math. Annalen*, t. 21 (1883), p. 545-591.
- CANTOR (XXIV). — *Bemerkung mit Bezug auf den Aufsatz: zur Weierstrass-Cantorschen Theorie der Irrationalenzahlen. Math. Ann.*, t. 33 (1889), p. 476.
- DEDEKIND (II). — *Stetigkeit und irrationale Zahlen*, 1^{re} édit. Braunschweig Vieweg, 1888, 6^e éd., 1930.
- GRASSMANN (I). — *Lehrbuch der Arithmetik*, 1861 (Sämmt. W., II, 1, p. 295-349).
- HEINE (III). — *Die Elemente der Functionenlehre. Journal f. reine u. angew. Mathematik*, t. 74 (1872), p. 172.
- HELMHOLTZ (I). — *Zählen u. Messen erkenntnisstheoretisch betrachtet. Phil. Aufsätze E. Zeller gewidm.*, 1887, p. 17-52.
- ILLIGENS (I). — *Zur Weierstrass-Cantorschen Theorie d. irrationalen Zahlen Math. Annalen*, t. 33 (1889), p. 155-166.
- (II). — *Über die Definition der irrationalen Zahlen*, *ibid.*, t. 35.
- KOSSAK (I). — *Die Elemente der Arithmetik*. Programm des Werderschen Gymnasiums, 1872, p. 1-29.
- LIPSCHITZ (I). — *Grundlage der Analysis*, Bonn, 1877.
- MÉRAY (I). — *Nouveau Précis d'analyse infinitésimale*. Paris, Savy, 1872.
- PASCH (I). — *Einleitung in die Differential u. Integralrechnung*. Leipzig, 1882.
- (II). — *Über die Einführung der irrationalen Zahlen. Math. Annalen*, t. 40 (1892), p. 149-152.
- TANNERY J. (I). — *Introduction à l'étude des fonctions d'une variable*. Paris, Hermann, 1886.
- Pour une bibliographie complète sur les nombres irrationnels. V. Pringsheim-Molk. *Nombres irrationnels in Encyclopedie des sciences mathématiques*, I, 3.

§ 3. — *Séries trigonométriques.*

A. — *Ouvrages généraux et histoires.*

- HANKEL (I). — *Untersuchungen über die unendlich oft oszillierenden u. unstetigen Functionen* Gratulationsprogramm d. Tübinger Universität (1870), reprod. in *Math. Annalen*, t. 20 (1882), p. 63-112.
- LEBESGUE (I). — *Leçons sur l'intégration et la recherche des fonctions primitives*, 1^{re} éd. Paris, Gauthier-Villars, 1903, 2^e éd. augmentée, 1928.
- (II). — *Leçons sur les séries trigonométriques*. Paris, Gauthier-Villars, 1905.
- RIEMANN (I). — *Über die Darstellbarkeit einer willkürlichen Function durch eine trigonometrische Reihe. Abhandl. Götting Gesellsch.*, t. 13 (Habilitationsschrift 1854). *Math. Werke*, p. 228.

SACHSE A. (I). — *Versuch einer Geschichte der Darstellung willkürlicher Functionen einer Variabeln durch trigonometrische Reihen*, Inaugural Dissertation Göttingen 1879, trad. in *Bulletin d. Scs mathématiques*, t. 4 (1880), p. 43-64 et 83-112.

B. — Les travaux autour de Dirichlet et de Riemann.

ASCOLI (I). — *Sulla serie di Fourier*. *Annali di Math.* (II) t. 4 (1873).

Du Bois REYMOND P. (V). — *Über die sprungweise Werthveränderungen analytischer Functionen*. *Math. Ann.*, t. 7 (1873), p. 241.

— (V). — *Über die Darstellung stetiger Functionen durch trigonometrische Reihen*. *Götting Nachrichten*, t. 1 (1873), p. 572-574.

— (X). — *Beweis dass die Coefficienten der trigonometrischen Reihen...* *Abhdl. d. bayer. Akad. Math. Nat. Wiss Kl.* t. XII, 1875), p. 117-167.

— (XI). — *Untersuchungen über die Convergenz u. Divergenz der Fourierschen Darstellungsformeln*, *ibid.*, IIte Abteilung (1876), p. 1.

DUHAMEL (I). — *Note sur la discontinuité des séries et sur les moyens de la reconnaître*. *Journ. Louville*, t. 12 (1854) p. 112-120.

HEINE E. (I). — *Handbuch d. Kugelfunctionen*. Berlin, 1861.

— (II). — *Über die trigonometrische Reihen*. *Journ f. reine u. angew. Math.*, t. 71 (1870), p. 352.

GENOCCHI A. (I). — *Interno ad alcune serie*. *Atti d. reale Acc. d. Sc. d. Torino*, t. 10 (1875).

LEJEUNE DIRICHLET (I). — *Sur la convergence des séries trigonométriques qui servent à représenter une fonction arbitraire entre des limites données*. *Journ. f. reine u. angew. Math.*, t. 4 (1829), p. 157.

— (II). — *Über die Darstellung willkürlicher Functionen durch trigonometrischen Reihen*. *Dove's Repertorium*. I (1837).

— (III). — *Sur les séries dont le terme général dépend de deux angles*. *Journ. f. r. u. angew. Math.* t. 17 (1837).

— (IV). — *Sur l'usage des intégrales définies dans la sommation des séries finies ou infinies*, *ibid.*

LIPSCHITZ (I). — *De explicatione per series trigonometricas instituenda functionum unius variabilis arbitrarium et praecipue earum quae per variabilis spatium finitum valorum maximorum et minimorum numerum habent infinitum disquisitio*. *Journ. f. r. u. angew. Math.* t. 63 (1864) p. 296 — trad. française av. notes par P. Montel in *Acta mathematica* t. 36 (1913) p. 281 (les notes du texte renvoient à la trad. Montel).

SCHÄFFLI (I). — *Einige Zweifel an d. allgemeinen Darstellbarkeit einer willkürlicher periodischen Function einer variablen durch eine trigonometr. Reihe*, Berne Univ. Programme, 1874.

— (II). — *Über die partielle Differential Gleichungen* $\frac{\partial \omega}{\partial t} = \frac{\partial^2 \omega}{\partial x^2}$. *Journ. f. r. u. angew. Math.*, t. 72 (1870), p. 283.

C. — Unicité du développement et notion de convergence uniforme.

- ABEL (I). — *Untersuchungen über die Reihe* $1 + \frac{m}{1} x + \frac{m(m-1)}{1-2} x^2 + \dots$
Journ. f. r. u. angew. Math., t. I (1826), p. 311.
- CANTOR G. (I). — *Über einen die trigonometrische Reihen betreffenden Lehrsatz*. *Journ. f. r. u. angew. Math.*, t. 72 (1870), p. 130-8.
- (II). — *Beweis dass eine für jeden reellen Wert von x durch eine trigonometrische Reihe gegebene Funktion $f(x)$ sich nur auf eine einzige Weise in dieser Form darstellen lässt.*, *ibid.*, t. 72 (1870), p. 139-142.
- (III). — *Notiz zu dem Aufsätze: Beweis dass eine...*, *ibid.*, t. 73 (1871), p. 292-294.
- (IV). — *Über trigonometrische Reihen*. *Math. Annalen*, t. 4 (1871), p. 139-143.
- (V). — *Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen*. *Math. Annalen*, t. 5 (1872), p. 123-132.
- CAUCHY (I). — *Cours d'analyse algébrique*. Paris, 1821.
- (II). — *Notes sur les séries convergentes dont les divers termes sont des fonctions continues d'une variable réelle ou imaginaire entre les limites données*. *C. R. Acad. des Sciences*, t. 36 (1853), p. 454-459.
- DINI U. (I). — *Sulla serie die Fourier*. *Annali di Matematica* (II), t. 6 (1873), p. 21 et 298.
- HEINE E. (III). — *Die Elemente der Functionenlehre*. *Journ. f. r. u. angew. Math.*, t. 74 (1872), p. 172.
- SCHWARZ (I). — *Beweis eines für die Theorie der trigonometrischen Reihen in Betracht kommenden Hilfsatzes*. *Ges. Abhandl.* Berlin, 1890, t. II, p. 341-343.
- SEIDEL (I). — *Note über eine Eigenschaft von Reihen welche discontinuirliche Functionen darstellen*. *Abh. d. bayer. Akad.*, 1847-1849.
- THOMÉ (I). — *Über hypergeometrische Kettenbrüche*. *Journ. f. r. u. angew. Math.*, t. 66, p. 334

et les ouvrages cités en A) et B).

§ 4. — *Théorie de la croissance.*

- BOREL E. (I). — *Leçons sur la théorie des fonctions*, 1^{re} édit., Paris. Gauthier-Villars, 1898 ; 3^e éd. augm., 1928. Note II sur la théorie de la croissance — (les notes du texte renvoient à la 3^e éd.).
- (II). — *Leçons sur les fonctions entières*, 1^{re} édit. Paris. Gauthier-Villars, 1900 ; 2^e éd. 1920.
- (III). — *Leçons sur les séries à termes positifs*, Paris. Gauthier-Villars, 1902.
- (IV). — *Leçons sur la théorie de la croissance*. Paris. Gauthier-Villars, 1909.
- DU BOIS REYMOND P. (I). — *Bemerkungen über die verschiedenen Werthe*

- welche eine Function zweier reellen Variabeln erhält, wenn man diese Variabeln entweder nacheinander oder Gewissen Beziehungen gemäss gleichzeitig verschwinden lässt. *Journ. f. r. u. angew. Math.*, t. 70 (1868), p. 10-45.
- (II). — *Sur les infinis des fonctions.* *Annali di matematica*, t. 4 (1871).
 - (III). — *Théorème général concernant la grandeur relative des infinis des fonctions et de leurs dérivées.* *Journ. f. r. u. angew. Math.*, t. 74 (1871).
 - (IV). — *Neue Theorie der Convergenz u. Divergenz von Reihen mit positiven Gliedern*, *ibid.*, t. 76 (1872), p. 60.
 - (VII). — *Über asymptotische Werthe, infinitäre Approximationen und infinitäre Auflösung von Gleichungen.* *Math. Ann.*, t. 8 (1874), p. 362.
 - (VIII). — *Zwei Sätze über Grenzwerte von Functionen zweier Veränderlichen*, *ibid.*, t. 11 (1877), 145-8.
 - (IX). — *Über die Paradoxen des Infinitärcalculs*, *ibid.*, t. 11 (1877), p. 149.
 - (XII). — *Über Integration und Differentiation infinitärer Relationen* *Math. Ann.*, t. 14, p. 498-501.
 - (XIII). — *Allgemeine Funktionen Theorie* Tübingen (1882).
- HAHN H. (I). — *Über die nichtarchimedischen Grössensysteme.* *Sitz. Ber. d. kaiser. Akad. d. Wiss. Math. Nat. Wiss. Kl.*, t. 116, II a (1907), p. 601.
- HARDY G. H. (I). — *Orders of Infinity.* London, 1910.
- PRINGSHEIM (I). — *Über die sogenannte Grenze u. die Grenzgebiete zwischen Convergenz u. Divergenz.* *Sitz. Ber. d. bayer. Akad. Math. Nat. Wiss. Kl.*, t. 26 (1896).
- (II). — *Über die Du Bois Reymond'sche ConvergenzGrenze und eine besondere Form der Convergenz Bedingung für unendliche Reihen*, *ibid.*, t. 27 (1897).
 - (III). — *Allgemeine Theorie d. Convergenz u. Divergenz.* *Math. Ann.*, t. 35 (1890) et 39 (1891).
- STOLZ O. (I). — *Über die Grenzwerte der Quotienten.* *Math. Ann.*, t. 14 (1878), p. 231-240.
- (II). — *Zur Geometrie der Alten, insbesondere über ein Axiom des Archimedes.* *Math. Ann.*, t. 22 (1883), p. 504.

CHAPITRE II

§ 1. — *Bolzano et Dedekind.*

- BOLZANO (II). — *Paradoxien des Unendlichen* herausggben von Prihonsky, Berlin, Mayer et Müller, 1899.
- CANTOR (XV). — Voir § suivant.
- DEDEKIND (I). — *Allgemeine Sätze über Räume* inédit publié par E. Noether in Dedekind (V), t. II, p. 353-355.

§§ 2-5. — *La première création cantorienne.*

- BENDIXSON (I). — *Quelques théorèmes de la théorie des ensembles de points* *Acta mathematica*, t. II (1883), p. 414.
- CANTOR (VI). — *Über eine Eigenschaft des Inbegriffs aller reellen algebraischen Zahlen.* *Journ. f. r. u. angew. Math.*, t. 77 (1874), p. 258-263.
- (VII). — *Ein Beitrag zur Mannigfaltigkeitslehre*, *ibid.*, t. 84 (1877), p. 242.
- (VIII). — *Über einen Satz aus der Theorie der stetigen Mannigfaltigkeiten.* *Götting. Nachrichten*, 1879, p. 127-135.
- (IX). — *Über unendliche lineare Punktmannigfaltigkeiten*, I, *Math. Annalen*, t. 15 (1870), p. 1-18.
- (X). — *Über unendliche lineare Punktmannigfaltigkeiten*, II, *ibid.*, t. 17 (1880), p. 355-359.
- (XI). — *Bemerkungen über trigonometrische Reihen.* *Math. Ann.*, t. 16 (1880), p. 113-115; *Fernere Bemerkungen über trigonometrische Reihen*, *ibid.*, p. 267-269.
- (XII). — *Über ein neues u. allgemeines Kondensationsprinzip der Singularitäten von Funktionen.* *Math. Ann.*, t. 19 (1882), p. 588-594.
- (XIII). — *Über unendliche lineare Punktmannigfaltigkeiten*, III. *Math. Ann.*, t. 20 (1882), p. 183.
- (XIV). — *Über unendliche lineare Punktmannigfaltigkeiten*, IV, *ibid.*, (1883), p. 51.
- (XV). — *Über unendliche lineare Punktmannigfaltigkeiten*, V. *ibid.*, p. 545-586.
- (XVI). — *Sur divers théorèmes de la théorie des ensembles de points dans un espace continu à n dimensions.* *Acta Math.*, t. II (1883), p. 409-414.
- (XVII). — *Gundlagen einer allgemeinen Mannigfaltigkeitslehre*, Leipzig Teubner, 1883 (réunion des mémoires XI, X, XII, XIV, XV).
- (XVIII). — *Über unendliche lineare Punktmannigfaltigkeiten*, VI, *Math. Ann.*, t. 23 (1884), p. 453-488.
- (XIX). — *De la puissance des ensembles parfaits de points.* *Acta Math.*, t. 4 (1884), p. 381-392.
- (XXIII). — *Mitteilungen zur Lehre von Transfiniten*, I et II, *Zeitschr. f. Philos. u. phil. Kritik*, t. 91 (1887) p. 81-125 et 252-270 et t. 92 (1888), p. 240-265.
- (XXV). — *Über eine elementare Frage der Mannigfaltigkeitslehre* *Jahresber. d. deutsch. Math. Verein.*, t. 1 (1892), p. 75-78.
- DU BOIS REYMOND (I). — *Allgemeine Funktionentheorie* (déjà cité).
- GRASSMANN (II). — *Ausdehnungslehre*, 2^e éd. 1862.
- JOURDAIN (I). — *The development of the theory of transfinite numbers*, I, *Arch. d. Math. u. Physik* (3) : I, t. 10 (1905), p. 244-281; II, t. 14 (1909) p. 289-311; III : t. 16 (1910) p. 21-43; IV : t. 22 (1914) p. 1-21.
- SCHENFLIES (II). — *Die Krisis in Cantors mathematischen Schaffen.* *Acta math.*, t. 50 (1928), p. 1, sq.

§ 6. — *Systématisation de la théorie.*

- BOREL E. (I). — *Leçons sur la théorie des fonctions* (cité pour ch. I, § 4).
- CANTOR (XX). — *Zum Problem des aktualen unendlichen*. Zeitschr. f. Philos. u. phil. Kritik, t. 88 (1885) p. 224-233.
- (XXI). — *Über verschiedene Theoreme aus der Theorie der Punktmen- gen in einem n-fach ausgedehnten stetigen Raume G_n* . Acta Math., t. 7 (1885), p. 105-124.
- (XXII). — *Über die verschiedenen Ansichten in Bezug auf die aktual unendlichen Zahlen*. Stockholm Bihang till kongl. svenska Vetenskaps Akad. Handlingar, t. 11 (1886) reproduction avec quelques modi- fications de (XX).
- (XXIII) *Mitteilungen zur Lehre von Transfiniten I et II* (déjà cité in § 2-5).
- (XXVI) *Beiträge zur Begründung der transfiniten Mengenlehre I*. Math. Ann. t. 46 (1895) p. 481-512.
- (XXVII). — *Beiträge zur Begründung der transfiniten Mengenlehre II*, ibid., t. 49 (1897), p. 343-347.
- (XXVI) et (XXVII) trad. en français par Marotte in *Mémoires de la Société des sciences physiques et naturelles de Bordeaux* (5) t. 3 (1899), p. 343-437, trad. en anglais par Ph. E. B. Jourdain, Chicago 1915.
- (XXIX). — *Sui numeri transfiniti* (Estratto d'una lettera di Georg Cantor a G. Vivanti) Rivista di Matemat., t. 5 (1895), p. 104-108.
- (XXX). — *Lettera di Georg Cantor a G. Peano*, ibid., p. 108-109.
- DEDEKIND R. (IV). — *Zweite Definition des Endlichen und Unendlichen*, opusculé inachevé publié in Dedekind (V) t. III.
- DU BOIS REYMOND (IX). — *Über die Paradoxen des Infinitärkalküls* (cité pour ch. I, § 4).
- (XII). — *Über Integration u Differentiation infinitärer Relationen* (id).
- STOLZ (II). — *Zur Geometrie der Alten, insbesondere über ein Axiom des Archi- medes* (id).
- VERONESE (I). — *Fundamenti di Geometria*. Padua, 1891.

§ 7. — *Le paradoxe Burali-Forti.*

- BERNSTEIN F. (I). — *Über die Reihe der transfiniten Ordnungszahlen*. Math. Ann., t. 60 (1905), p. 187-193.
- BURALI-FORTI C. (I). — *Una questione sui numeri transfiniti*. Rendiconti del Circolo Matemat. di Palermo, t. 11 (1897), p. 154.

CHAPITRE III

§ 1. — *L'œuvre de Dedekind.*A. — *La création de 1887 et les ouvrages qui la préparent.*

- DEDEKIND R. (III). — *Was sind und was sollen die Zahlen ?* 1^{re} éd. Brauns- chweig Vieweg u. Sohn 1887, 3^e éd., 1911.
- HELMHOLTZ (I). — *Zählen und messen erkenntnis theoretisch betrachtet*, Leip- zig, 1887.

- KRONECKER (I). — *Über den Zahlbegriff*, Leipzig, 1887.
 LEJEUNE-DIRICHLET (V). — *Vorlesungen über Zahlentheorie*, 3^e éd. 1879.
 SCHRÖDER E. (I). — *Lehrbuch der Arithmetik und Algebra*, 1^{er} Band. Leipzig, 1873.
 ZERMELO (II). — V. ci-après, § 2-3.

B. — Les ensembles finis.

- CAVAILLÈS J. (I). — *Sur la 2^e définition des ensembles finis donnée par Dedekind. Fund. Mathematicae*, t. 19 (1932), p. 143.
 DEDEKIND R. (IV). — *Zweite Definition des Endlichen und Unendlichen*, opuscule inachevé publié in Dedekind (V).
 VON NEUMANN J. (IV). — *Die Axiomatisierung der Mengenlehre. Math. Zeitschr.*, t. 27 (1928), p. 669.
 TARSKI A. (II). — *Sur les ensembles finis. Fund. Mathematicae*, t. 6 (1925), p. 45.
 WEBER H. (I). — *Arithmetik, Algebra, Analysis* in Weber-Wellstein *Enzyklopädie der Elementarmathematik*, t. 1, 4^e éd. Leipzig et Berlin, 1922.
 WHITEHEAD and RUSSELL (I). — *Principia Mathematica*, Cambridge, 1^{re} éd., t. I (1910), t. II (1912), t. III (1913), 2^e éd., t. I (1925); t. II, III non modifiés (1927).
 ZERMELO E. (IV). — *Sur les ensembles finis et le principe d'induction complète Acta Mathematica*, t. 32 (1909), p. 185.

§ 2. — L'axiome de choix et le théorème du bon ordre.

- BERNSTEIN F. (II). — *Über den Beweis der Wohlordnung Math. Ann.*, t. 60 (1905), p. 463.
 — (III) *Untersuchungen aus der Mengenlehre*, *ibid.*, t. 61 (1905), p. 117-165 (Reproduction de : Inaugural Dissertation Göttingen, 1901).
 BOREL E. (II). — *A propos d'un théorème de M. Zermelo. Math. Ann.*, t. 60 (1905).
 FRAENKEL A. (III). — *Über den Begriff «definit» und die Unabhängigkeit des Auswahlaxioms. V. ci-après § 3-5.*
 — (XIII). — *Sur l'axiome du choix, L'Enseignement math.*, 1935, p. 32.
 — (XIV) *Über eine abgeschwächte Fassung des Auswahlaxioms. V. ci-après § 3-5.*
 KÖNIG J. (II). — *Über die Grundlagen der Mengenlehre und das Kontinuumproblem I : Math. Ann.*, t. 61 (1905), p. 156-160. II : *ibid.*, t. 63 (1907), p. 217, trad. in *Acta Mathematica*, t. 30 (1906), p. 329 et t. 31 (1908), p. 89.
 LÉVI B. (I). — *Intorno alla teoria degli aggregati. R. Istituto Lombardo di scienze e lettere Rendic.* (2), t. 35 (1902), p. 863.
 TARSKI (I). — *Sur quelques théorèmes qui équivalent à l'axiome de choix. Fund. Mathematicae*, t. 5 (1924), p. 47.
 ZERMELO E. (I). — *Beweis dass jede Menge wohlgeordnet werden kann. Math. Ann.*, t. 59 (1904), p. 514-516.
 — (II) *Neuer Beweis für die Wohlordnung*, *ibid.*, t. 65 (1908), p. 107-128.

§ 3-5. — *L'Antinomie Russell et l'axiomatisation Zermelo-Fraenkel.*

- BAER R. (I). — *Zur Axiomatik der Kardinalzahlarithmetik. Math. Zeitschr.* t. 29 (1929).
- FRAENKEL (I). — *Axiomatische Begründung der transfiniten Kardinalzahlen, I. Math. Zeitschrift*, t. 13 (1922), p. 153-188.
- (II). — *Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre Math. Annalen*, t. 86 (1922), p. 230-237.
- (III). — *Über den Begriff «definit» und die Unabhängigkeit des Auswahlaxioms. Sitzungsab. d. Preuss. Akad. d. Wiss. Phys.-math. Kl.* 1922, p. 253-257.
- (IV). — *Die neueren Ideen zur Grundlegung der Analysis und Mengenlehre. Jahresb. d. dtsh. Math. Verein*, t. 33 (1924), p. 97-103.
- (V). — *Untersuchungen über die Grundlagen der Mengenlehre Math. Zeitschrift* t. 22 (1925), p. 250-273.
- (VI). — *Axiomatische Theorie der Geordneten Mengen Journ. f. Mat.* t. 155 (1926), p. 129-158.
- (VII). — *Zehn Vorlesungen über die Grundlegung der Mengenlehre.* Leipzig u. Berlin Teubner 1927.
- (VIII). — *Über die Gleichheitsbeziehung in der Mengenlehre. Journ. f. Math.* t. 157 (1927), p. 79-81.
- (IX). — *Über die Ordnungsfähigkeit beliebiger Mengen. Sitzungsab. d. Preuss. Akad. d. Wiss. Phys. math. Kl.* 1928.
- (XII). — *Axiomatische Theorie der Wohlordnung. Journ. f. Math.*, t. 167 (1932), p. 1-11.
- (XIII). — *Sur l'axiome du choix. Conférence faite à Genève en 1934. L'enseignement mathématique*, 1935, p. 32-51.
- (XIV). — *Über eine abgeschwächte Fassung des Auswahlaxioms. The Journal of symbolic Logic*, t. 2 (1937), p. 1-25.
- HESSENBERG G. (I). — *Grundbegriffe der Mengenlehre* Sonderdruck aus den *Abhandl. d. Friesschen Schule* Bd 1 Heft 4 Göttingen, 1906.
- KURATOWSKI C. (I). — *Sur la notion de l'ordre dans la théorie des ensembles. Fund. Mathematicae* t. 2 (1921), p. 161-171.
- MAHLO P. (I). — *Über lineare transfinite Mengen. Ber. über d. Verhandl. d. K. Sächs. Ges. d. Wiss. 7. Leipzig Math. phys. Kl. t. 63* (1911), p. 187-225.
- (II). — *Zur Theorie und Anwendung der $\aleph_0$ -Zahlen*, *ibid.* t. 64 (1912), p. 108-112 et t. 65 (1913), p. 268-282.
- VON NEUMANN J. (V). — *Über die Definition durch transfinite Induktion und verwandte Fragen der allgemeinen Mengenlehre avec un complément de A. Fraenkel. Math. Ann.* t. 99 (1928), p. 373-393.
- RUSSELL B. (I). — *The principles of mathematics* I Cambridge 1903.
- ZERMELO (I). — *Beweis dass jede Menge wohlgeordnet werden kann. Math. Annalen*, t. 59 (1904), p. 514-516.
- (II). — *Neuer Beweis für die Wohlordnung*, *ibid.*, t. 65 (1908), p. 107-128.
- (III). — *Untersuchungen über die Grundlagen der Mengenlehre I*, *ibid.* t. 65 (1908), p. 261-281.

§ 5. — *Les ordinaux transfinis.*

- HESSENBERG G. (II) *Kettentheorie und Wohlordnung.* Journ. f. r. u. angew. Math. t. 135 (1909), p. 81-133.
- VON NEUMANN J. (I). — *Zur Einführung der transfiniten Ordnungszahlen.* Acta Univ. Franz. Jos. Szeged 1 (1923), p. 199-208.
- (II). — *Eine Axiomatisierung der Mengenlehre.* Journ. f. r. u. angew. Math., t. 154 (1925), p. 219-240.
- (III). — *Die Axiomatisierung der Mengenlehre.* Math. Zeitschr., t. 27 (1928), p. 669-752.
- (IV). — *Über die Definition durch transfinite Induktion etc.* Math. Ann., t. 99 (1928), p. 373.
- (V). — *Die Widerspruchsfreiheit eines Axiomensystems der Mengenlehre.* Journ. f. r. u. angew. Math. t. 160.

§ 6. — *L'absence de catégoricité, le paradoxe de Skolem.*

- EKLUND H. (I). — *Über Mengen die Elemente ihrer selbst sind.* Nyt Tidsskr. f. Mathematik. Avd. 13, t. 29 (1918), p. 8-28.
- MIRIMANOFF D. (I). — *Les antinomies de Russell et de Burali-Forti et le problème fondamental de la théorie des ensembles.* L'enseignement math., t. 19 (1917), p. 37-52.
- (II). — *Remarques sur la théorie des ensembles et les antinomies cantorienes*, I, ibid, t. 19 (1917), p. 209-217 ; II, ibid, t. 21 (1920), p. 29-52.
- VON NEUMANN (II) et (V).
- SKOLEM Th. (I). — *Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze.* Skr. utg. av. Videnskaps. Kristiania 1919, I, Math. Naturv. Kl., p. 1-37.
- (II). — *Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre.* 5 Kongress d. skandinav. Math. Helsingfors, 1922, p. 217.
- (III) *Über einige Grundlagenfragen der Mathematik.* Skr. Nor. Vid. Akad. in Oslo, Mat-Naturv. Kl., 1929, n° 4.

CONCLUSION

- HARTOGS F. (I). — *Über das Problem der Wohlordnung.* Math. Ann., t. 76 (1915), p. 438-443.
- KÖNIG J. (I). — *Zum Kontinuumproblem.* Math. Ann., t. 60 (1905), p. 177-180 et 462. 1^{er} redact. in Verh. d. 3. internat. Math. Kongress in Heidelberg, (1904).
- LUSIN N. (I). — *Sur les ensembles analytiques nuls.* Fund. Math., 185 (1935), p. 109-131.

INDEX DES TERMES EMPLOYÉS (1)

A

Aleph (puissance d'un ensemble bien ordonné) : 81, 94.
Apantachie : 39.
Application : 51 n.
Axiomatique (méthode) : 18.
Axiome-N : 133.

B

Bien ordonné (ensemble) : 64.
Biunivoque (correspondance) : 46.
Bon ordre : 64.

C

Cardinal (nombre) : 75.
Catégoricité : 5 n.
Chaîne : 100.
Chaîne propre (d'un ensemble) : 101.
0-Chaîne : 117.
Choix (axiome de) : 113.
Classes I, II, III des nombres ordinaux transfinis : 67.
Compacité en soi : 52.
Comparabilité des puissances : 66.
Consistante (multiplicité) : 95.
Continuité : 16.
Coupure : 16.
Critiques (nombres) : 91.
Croissance (ordre de) : 40.
Cyclique (ordre) : 111.

D

Définie (propriété) (Zermelo) : 123.
Dénombrable (ensemble) : qui peut être mis en correspondance biunivoque avec l'ensemble des entiers : 50.
Dense (ensemble) dans un intervalle : 58.
Dense en soi (ensemble) : 69, 85.
Dense (partout non) : 58.

Densité : 52.

Dérivation (d'un ensemble) : 22.

Diagonale (procédé de la) : 42, 72.

Distributive (procédé de la propriété) : 12, 70.

E

Emboîtement des intervalles (procédé de l') : 52.

Equivalence (théorème de l') : 78, 109 n.

Equivalents (ensembles) : 55, 75.

Espèce (d'un ensemble de points) : 22.

Espèce d'un nombre ordinal (1^{re} ou 2^e espèce) : 88.

Exponentiation d'un ensemble par un autre : 77.

F

Fermé (ensemble) : 71, 85.

Fusionnement (procédé de) : 106.

G

Genre (d'un ensemble) : 58.

Groupe intégrable : 58.

I

Intersection de 2 ou un nombre quelconque d'ensembles : 100 n.

Infinitaire (calcul) : 40.

— (ordre) : 41.

Irréductible (élément) d'un ensemble : 113.

Isolé (ensemble) : 60.

L

Linéarisation : 51.

N

Nombrabilité (d'un ensemble) : 131

Nombrement : 131.

(1) Les pages indiquées sont celles où le terme est défini ou (en l'absence d'une définition explicite, qui est alors donnée ici) utilisé pour la première fois.

Nombres-ε : 92.

Normales (Fonctions) : 91.

O

Ordinal (nombre) : 64.

P

$\mathfrak{P}M$: ensemble des parties de l'ensemble M : 119.

Pantachie : 39.

Parfait (ensemble) : 69, 85.

Partie d'un ensemble : 55.

Principal (élément), d'un ensemble ordonné : 84.

Principes d'engendrement, d'arrêt : 65, 67.

Produit d'ensembles : 77.

Propriété distributive ou propriété Υ (procédé de la) : 12, 70.

Puissance d'un ensemble : 55, 75.

Q

Quasi-Chaine : 111.

R

Réurrence (définition par) : finie : 105.

Réductible (ensemble) : dont le *vième* dérivé ($\vee$ étant un nombre fini ou un ordinal transfini de la classe II) est fini : 68.

Reste d'un ensemble ordonné (par rapport à un segment) : 85.

Réunion (ensemble) : 77.

S

$\mathfrak{S}M$ (ensemble-réunion des ensembles éléments de M) : 122.

Saturé (élément) d'un ensemble : 113.

Segment, dans un ensemble ordonné : 85.

Semblable (application) : 82.

Séparation (axiome de) : 122.

Sous ensemble : V. partie d'ensemble.

Subsorption (ordre de) : 130.

Substitution (axiome de) : 127.

T

Type d'ordre : 81.

TABLE DES MATIÈRES

	Pages
INTRODUCTION	5
CHAPITRE I. — <i>Préhistoire : Nombres irrationnels, séries trigonométriques, théorie de la croissance</i>	10
§ 1. La démonstration Bolzano du théorème fondamental de l'analyse.....	11
§ 2. Les définitions des nombres irrationnels.....	14
a) définition Dedekind.....	15
b) définition Weierstrass.....	18
c) définition Heine-Cantor.....	20
§ 3. Les séries trigonométriques.....	23
a) la notion de fonction arbitraire : origine du problème... ..	23
b) les travaux de Dirichlet-Lipschitz.....	26
c) le mémoire de Riemann :.....	32
1. nouvelle définition de l'intégrale.....	32
2. les conditions nécessaires pour la représentation.....	34
d) le problème de l'unicité du développement d'une fonction :	
1. convergence uniforme et intégration terme à terme..	35
2. travaux de Heine et de Cantor.....	37
§ 4. Théorie de la croissance de Du Bois Reymond.....	40
a) idée du continu infinitaire.....	40
b) première utilisation du procédé de la diagonale.....	42
CHAPITRE II. — <i>La création cantorienne</i>	45
§ 1. Pressentiments d'une théorie des ensembles.....	45
a) les <i>Paradoxes de l'Infini</i> de Bolzano.....	45
b) les <i>Propositions générales sur les espaces</i> de Dedekind... ..	49
§ 2. La découverte des deux puissances.....	50
a) puissance de l'ensemble des nombres réels, distinguée de celle de l'ensemble des nombres entiers.....	51
b) invariance de la puissance dans les continus à un nombre quelconque de dimensions.....	52

§ 3. L'opération de la dérivation et la mise en rapport des deux puissances	57
§ 4. La création de 1882-1884.....	61
a) les ordinaux transfinis.....	63
b) ensembles fermés et ensembles parfaits.....	69
§ 5. Le procédé général de la diagonale.....	72
§ 6. Systématisation de la théorie.....	74
a) théorie cardinale.....	75
b) les ensembles finis.....	79
c) les types d'ordre.....	81
d) théorie ordinale :.....	85
1. les ensembles bien ordonnés.....	85
2. la récurrence transfinie.....	88
3. forme canonique des nombres ordinaux.....	90
4. les nombres critiques ϵ	92
§ 7. Le paradoxe de Burali-Forti : essai d'une solution au problème du bon ordre (les multiplicités inconsistantes).....	94
CHAPITRE III. — <i>Dedekind et la chaîne — les axiomatisations</i>	97
§ 1. Préparation d'une théorie abstraite par Dedekind.....	97
a) définition du matériel de la théorie générale. — La chaîne.....	99
b) Les ensembles <i>simplement infinis</i> :.....	102
1. théorème d'existence de l'infini.....	103
2. la suite des entiers et l'ordre.....	104
3. fondement des définitions par récurrence.....	105
c) les ensembles finis :.....	110
1. première définition (<i>Was sind u. was sollen die Zahlen ?</i>).....	110
2. l'ordre cyclique.....	111
§ 2. Les deux démonstrations du bon ordre, grâce à la chaîne : l'axiome de choix.....	113
§ 3. L'axiomatisation de Zermelo.....	118
a) l'antinomie Russell.....	119
b) les axiomes de Zermelo.....	121
§ 4. Elaboration critique (Fraenkel).....	124
a) précision du domaine de la théorie : l'axiome de séparation.....	124
b) l'indépendance de l'axiome de choix.....	128
§ 5. Les ordinaux transfinis (von Neumann).....	130

a) théorie générale. Bon ordre et nombrement.....	131
b) fondement des définitions par récurrence transfinie.....	132
c) l'axiomatique von Neumann ; nouvelle démonstration du bon ordre (par le paradoxe Burali-Forti).....	133
§ 6. Le paradoxe de Skolem et le manque de catégoricité dans toute axiomatisation de la théorie.....	135
CONCLUSION.....	138
BIBLIOGRAPHIE.....	142
INDEX DES TERMES EMPLOYÉS.....	152

